INTRODUCTION

The continuous and growing change and globalization are plunging individuals into a system that requires knowledge, skills and figures in the plot are following hierarchical clustering disposition. Finally, social network analysis is conducted, both with Complete Method and Density-based method (DBSCAN), obtaining three clusters provided in a dendrogram visualization (Fig. 3) and in a spatial bidimensional map. These results have led to an increasing interest in the analysis of the skills required for Data Scientists (DSs) in the professional market. The main objective of this research is to analyze the skills required by DSs and to compare them with the skills extracted from job advertisements.

BACKGROUND OVERVIEW

Nowadays, the increasing attention to the continuous amount of unstructured, semi-structured and structured data, obtained from websites and social networks requires the age of Big Data professionals for the companies to understand the market business context (Provost & Fawcett, 2013). In the literature, there are many case studies of the use of Data Scientists (DSs) in the software market development in the information technology sector. Usually, DSs are employed mainly in mining data and analyze target data to provide diagnostic, predictive and descriptive intelligence and business value. They are responsible for developing and implementing information systems, business intelligence and data mining, to understand data and also data vector (Holm, 2016). In the same context, the network presented a directed density equal to 0.518, showing a medium effective mapping. Mapping the main skills required by a professional figure allows companies to conduct independent research and analyses on enormous volumes of unstructured, semi-structured and structured data, obtained from websites and social networks.

RESULTS

The fourth and last phase of the methodology refers to the visualization of the skills through a network (Weinberg, 1962; Strogatz, 1998). The network presented a directed density equal to 0.518, showing a medium effective mapping. Mapping the main skills required by a professional figure allows companies to conduct independent research and analyses on enormous volumes of unstructured, semi-structured and structured data, obtained from websites and social networks.

data management software developer
machine learning
problem solving
computer science
work-in-team
soft skills
system sets
algorithm
artificial intelligence
big data management
analytics capability
business strategy alignment.

The fourth and last phase of the methodology refers to the visualization of the skills through a network (Weinberg, 1962; Strogatz, 1998) that highlights the main interconnections between the clusters derived from the Complete Method and Density-based method (DBSCAN), obtaining three clusters provided in a dendrogram visualization (Fig. 3) and in a spatial bidimensional map. These results have led to an increasing interest in the analysis of the skills required for Data Scientists (DSs) in the professional market. The main objective of this research is to analyze the skills required by DSs and to compare them with the skills extracted from job advertisements.

The sample was composed of 395 job ads collected from the Indeed portal in the American (US) market, recording 3.8 million job ads in 300,000 documents, including several DS job offers. The training phase was conducted with 395 job ads using the software DBSCAN (Ester et al., 1996), obtaining a spatial bidimensional map for the visualization. The network presented a directed density equal to 0.518, showing a medium effective mapping. Mapping the main skills required by a professional figure allows companies to conduct independent research and analyses on enormous volumes of unstructured, semi-structured and structured data, obtained from websites and social networks.

The fourth and last phase of the methodology refers to the visualization of the skills through a network (Weinberg, 1962; Strogatz, 1998) that highlights the main interconnections between the clusters derived from the Complete Method and Density-based method (DBSCAN), obtaining three clusters provided in a dendrogram visualization (Fig. 3) and in a spatial bidimensional map. These results have led to an increasing interest in the analysis of the skills required for Data Scientists (DSs) in the professional market. The main objective of this research is to analyze the skills required by DSs and to compare them with the skills extracted from job advertisements. The fourth and last phase of the methodology refers to the visualization of the skills through a network (Weinberg, 1962; Strogatz, 1998) that highlights the main interconnections between the clusters derived from the Complete Method and Density-based method (DBSCAN), obtaining three clusters provided in a dendrogram visualization (Fig. 3) and in a spatial bidimensional map. These results have led to an increasing interest in the analysis of the skills required for Data Scientists (DSs) in the professional market. The main objective of this research is to analyze the skills required by DSs and to compare them with the skills extracted from job advertisements.

The sample was composed of 395 job ads collected from the Indeed portal in the American (US) market, recording 3.8 million job ads in 300,000 documents, including several DS job offers. The training phase was conducted with 395 job ads using the software DBSCAN (Ester et al., 1996), obtaining a spatial bidimensional map for the visualization. The network presented a directed density equal to 0.518, showing a medium effective mapping. Mapping the main skills required by a professional figure allows companies to conduct independent research and analyses on enormous volumes of unstructured, semi-structured and structured data, obtained from websites and social networks.

The fourth and last phase of the methodology refers to the visualization of the skills through a network (Weinberg, 1962; Strogatz, 1998) that highlights the main interconnections between the clusters derived from the Complete Method and Density-based method (DBSCAN), obtaining three clusters provided in a dendrogram visualization (Fig. 3) and in a spatial bidimensional map. These results have led to an increasing interest in the analysis of the skills required for Data Scientists (DSs) in the professional market. The main objective of this research is to analyze the skills required by DSs and to compare them with the skills extracted from job advertisements. The fourth and last phase of the methodology refers to the visualization of the skills through a network (Weinberg, 1962; Strogatz, 1998) that highlights the main interconnections between the clusters derived from the Complete Method and Density-based method (DBSCAN), obtaining three clusters provided in a dendrogram visualization (Fig. 3) and in a spatial bidimensional map. These results have led to an increasing interest in the analysis of the skills required for Data Scientists (DSs) in the professional market. The main objective of this research is to analyze the skills required by DSs and to compare them with the skills extracted from job advertisements.