

Data Science Tools Application for Business Processes Modelling in Aviation

Maryna Nehrey

National University of Life and Environmental Sciences of Ukraine, Ukraine

Taras Hnot

National University of Life and Environmental Science of Ukraine, Ukraine

EXECUTIVE SUMMARY

Successful business involves making decisions under uncertainty using a lot of information. Modern modeling approaches based on data science algorithms are a necessity for the effective management of business processes in aviation. Data science involves principles, processes, and techniques for understanding business processes through the analysis of data. The main goal of this chapter is to improve decision making using data science algorithms. There are sets of frequently used algorithms described in the chapter: linear, logistic regression models, decision trees as a classical example of supervised learning, and k-means and hierarchical clustering as unsupervised learning. Application of data science algorithms gives an opportunity for deep analyses and understanding of business processes in aviation, gives structuring of problems, provides systematization of business processes. Business processes modeling, based on the data science algorithms, enables us to substantiate solutions and even automate the processes of business decision making.

INTRODUCTION

This chapter will address challenges with Data Science algorithms in aviation. Successful business involves making decisions under uncertainty using a lot of information. Modern modeling approaches based on Data Science algorithms are a necessity for the effective management of business processes in aviation. Data science involves principles, processes, and techniques for understanding business processes through the analysis of data. The main goal of this chapter is to improve decision making using data science algorithms. There are a set of frequently used algorithms described in the paper. Linear,

logistic regression models, decision trees as a classical example of supervised learning and k-means and hierarchical clustering as – unsupervised learning. Application of Data Science algorithms gives an opportunity for deep analyses and understanding of business processes in aviation, gives structuring of problems, provides systematization of business processes. Business processes modeling, based on the Data Science algorithms, enables us to substantiate solutions and even automate the processes of business decision making.

Nowadays business environment of aviation is characterized by significant uncertainty, increasing competition, and globalization. For a successful business operation, it is necessary to make decisions, taking into account a large number of factors and a considerable volume of information. The effectiveness of business decisions depends on the ability to analyze existing information, to predict the development of business processes and the system vision of the whole business. Business processes modeling in aviation is the most difficult part of their analysis. Improving the business decision-making process is possible provided that current methods and models of business analysis, such as Data Science, are correctly applied.

DATA SCIENCE: ESSENCE, PRINCIPLES, AND TOOLS

Data science involves principles, processes, and techniques for understanding business processes through the analysis of data. The main goal of Data Science is the exploration of the complexities inherently trapped in data, business, and problem-solving systems.

Data Science is a science of learning from data. Data Science is a continuation of Data Mining and Predictive Analytics. This approach is multidisciplinary, it combines the methods and models of disciplines such as mathematics, statistics, probability theory, information technology, including signal processing, probabilistic models, machine learning, statistical training, data mining, databases, object recognition, visualization, uncertainty modeling, data warehousing, data compression, computer programming, and high-performance computing.

The essence of Data Science is the extraction of information based on the knowledge and skills from the various fields of activity necessary for gaining knowledge. The composition of such a set largely depends on the field of research. For specialists in this area of research - Data Scientist - generalized qualification requirements have been developed.

Data Science has a big list of tools: Linear Regression, Logistic Regression, Density Estimation, Confidence Interval, Test of Hypotheses, Pattern Recognition, Clustering, Supervised Learning, Time Series, Decision Trees, Monte-Carlo Simulation, Naive Bayes, Principal Component Analysis, Neural Networks, k-means, Recommendation Engine, Collaborative Filtering, Association Rules, Scoring Engine, Segmentation, Predictive Modeling, Graphs, Deep Learning, Game Theory, Arbitrage, Cross-Validation, Model Fitting, etc. Some of this tools were used in the next researches. Teaching data science, for example, were introduced (Brunner & Kim, 2016), Big data and Data Science methods presented in (Chen, Chiang & Storey, 2012), (George, Osinga, Lavie & Scott, 2016), (Kucherov, 2007), (Shoro, Soomro, 2015), (Xiong, Yu & Zhang, 2017), machine learning used (Parish & Duraisamy, 2016), Monte Carlo method presented (Patriarca, Di Gravio & Costantino, 2017), game theory and genetic algorithms combined (Periaux, Chen, Mantel, Sefrioui & Sui, 2001), Artificial Intelligence presented (Rizun & Shmelova, 2017). Data Science is fast developing. A large volume of information that grows with each passing year makes it possible to build high-precision models that simplify and partially automate the

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/data-science-tools-application-for-business-processes-modelling-in-aviation/222188

Related Content

Quantization of Continuous Data for Pattern Based Rule Extraction

Andrew Hamilton-Wright and Daniel W. Stashuk (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1646-1652).

www.irma-international.org/chapter/quantization-continuous-data-pattern-based/11039

On Explanation-Oriented Data Mining

Yiyu Yao (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 842-848).

www.irma-international.org/chapter/explanation-oriented-data-mining/10918

Profit Mining

Senqiang Zhou (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1598-1602).

www.irma-international.org/chapter/profit-mining/11032

Formal Concept Analysis Based Clustering

Jamil M. Saquer (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 895-900).

www.irma-international.org/chapter/formal-concept-analysis-based-clustering/10926

Data Mining in Protein Identification by Tandem Mass Spectrometry

Haipeng Wang (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 472-478).

www.irma-international.org/chapter/data-mining-protein-identification-tandem/10862