Chapter 5
A Survey on Intelligence Tools for Data Analytics

Shatakshi Singh

Mody University of Science and Technology, India

Kanika Gautam

Mody University of Science and Technology, India

Prachi Singhal

Mody University of Science and Technology, India

Sunil Kumar Jangir

Mody University of Science and Technology, India

Manish Kumar

Mody University of Science and Technology, India

ABSTRACT

The recent development in artificial intelligence is quite astounding in this decade. Especially, machine learning is one of the core subareas of AI. Also, ML field is an incessantly growing along with evolution and becomes a rise in its demand and importance. It transmogrified the way data is extracted, analyzed, and interpreted. Computers are trained to get in a self-training mode so that when new data is fed they can learn, grow, change, and develop themselves without explicit programming. It helps to make useful predictions that can guide better decisions in a real-life situation without human interference. Selection of ML tool is always a challenging task, since choosing an appropriate tool can end up saving time as well as making it faster and easier to provide any solution. This chapter provides a classification of various machine learning tools on the following aspects: for non-programmers, for model deployment, for Computer vision, natural language processing, and audio for reinforcement learning and data mining.

INTRODUCTION

In this era of digitalization, data is generated and stored in every field. This data has the potential to be used to predict future possibilities. ML comprises of using this data to train a model for predicting results for new observations without any explicit instructions. It predicts results by analyzing the patterns in the given dataset.

Ranging from a very basic task of separating the spam emails to the revolutionary Cyborg technology ML has a wide range of applications in our life. It not only eases our tasks but has the potential to solve some of the most challenging issues of the society (Anirudhi Thanvi et al., 2019). A lot of research is going on to improve the quality and efficiency of medical care using ML. Digital diagnosis of life-threatening diseases like cancer helps to detect the problem at a very early stage, making it possible to treat it before it turns lethal. ML has great potential across several other domains like Education (Kotsiantis, 2012), Farming (Li et al., 2014; Paul et al., 2016), Environment Conservation (Rahman et al., 2014), Security (Buczak & Guven, 2016), Social Media (Alatabi & Abbas, 2020; ReddyPY.Sushmitha & M.Padma, 2016), Information Verification (Herrera et al., 2010; Sharma et al., 2019) and a lot more(Vyas & Jangir, 2019).

ML tools provide a more convenient way for developers to build, train, and deploy their models. Tools help to speed up the process by automating each step of the process which otherwise needed to be implemented manually from scratch. They help to eliminate the barrier between imagination and implementation of ideas due to a lack of skills. Developers need not be an expert in every step of implementation. They can implement their ideas easily once they are trained to work with a tool (Kumar Jangir et al., 2018, 2019).

The authors reviewed more than 50 papers and more than 100 blogs, articles, and websites that used various data analytics tools for carrying out different processes. The motivation behind carrying out this survey was the unavailability of research work that classified data analytics tools in various categories and gave a brief description of each tool.

This chapter divides data analytics tools into different categories for the simplicity and ease of choosing the right tool. It tells about the needs of tools for data analytics. The rest of the chapter is arranged as follows: Python Libraries used for carrying out various data analytics processes are explained followed by the tools for Big Data Analysis that help to organize and analyze large and complex datasets, which are difficult to be tackled by traditional database and software techniques. Then it states data mining tools that use mathematics, statistics, algorithms, and other scientific ways to provide insights from the given data set. The description is taken further by explaining tools that assist in Natural Language Processing, Voice recognition, face identification, and voice synthesizing which are classified as Tools for Computer Vision, NLP, and Audio. Then tools for Reinforcement learning that helps the agent to learn from its experiences are mentioned. After that tools that help to integrate the ML model with the existing production environment are categorized as tools for Model Deployment. At last, tools that obviate the need for programming skills are categorized under the category of tools for Non-Programmers. Various use cases of tools by different companies are given for each category. The use of some of the tools for research and development is also discussed along with the methodology, dataset description, and results used to carry out that research.

21 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/a-survey-on-intelligence-tools-for-data-analytics/264305

Related Content

Ethics in Predictive Learning Analytics: An Empirical Case Study on Students Perceptions in a Northern Irish University
Paul Joseph-Richardand James Onohuome Uhomoibhi (2021). *Advancing the Power of Learning Analytics and Big Data in Education (pp. 86-107).*
www.irma-international.org/chapter/ethics-in-predictive-learning-analytics/272948

Personal Diary Method: A Way of Collecting Qualitative Data
Farrah Zebaand Pankaj Kumar Mohanty (2019). *Qualitative Techniques for Workplace Data Analysis (pp. 96-116).*
www.irma-international.org/chapter/personal-diary-method/207793

Bayesian Kernel Methods: Applications in Medical Diagnosis Decision-Making Processes (A Case Study)
Arti Saxenaand Vijay Kumar (2021). *International Journal of Big Data and Analytics in Healthcare (pp. 26-39).*
www.irma-international.org/article/bayesian-kernel-methods/268416

Fitting a Three-Phase Discrete SIR Model to New Coronavirus Cases in New York State
Kris H. Green (2021). *International Journal of Data Analytics (pp. 59-74).*
www.irma-international.org/article/fitting-a-three-phase-discrete-sir-model-to-new-coronavirus-cases-in-new-york-state/285468

Simulating and Preventing COVID-19 Using Epidemiological Models
(2022). *Applied Big Data Analytics and Its Role in COVID-19 Research (pp. 28-57).*
www.irma-international.org/chapter/simulating-and-preventing-covid-19-using-epidemiological-models/300379