Chapter 3
Software Installation and Introduction to RapidMiner

ABSTRACT

Licensed and open-source data mining software has made it easier for data scientists to process and analyze large datasets. Python and R are popular open-source tools used for teaching data mining techniques, including data classification and clustering. However, one of the main challenges for domain experts looking to upskill in data science is computer programming. Licensed software such as RapidMiner provides a solution to this issue, as it is designed for data mining processing without requiring computer programming skills. RapidMiner allows data scientists to bring their data sets and mining models into an analysis process that includes all data science mechanisms. It also offers various supporting models for data pre-processing and creates data visualizations during data exploration. After data mining, RapidMiner displays easy-to-interpret results, enabling data scientists to effectively explain the answer to a business question to stakeholders. With its user-friendly interface, RapidMiner is an excellent choice for those without a technical background in programming.

INTRODUCTION

RapidMiner software is very popular in data science work, and has been used in data scientist development trainings (Chisholm, 2013: Klinkenberg & Hofmann, 2014: Kotu & Deshpande, 2014). The scientists can learn about various data mining techniques and apply those concepts in data analysis with RapidMiner without writing any program. Moreover, RapidMiner software supports every step of the data science field work. Data scientists can import datasets into RapidMiner software to promptly explore data through the feature visualizations. In data preparation, the scientists can use models within the RapidMiner software to customize their data ready for analysis with various data mining techniques. RapidMiner software provides the sample data sets for the scientists to use in experimenting data during their trainings as well.

RAPIDMINER SOFTWARE INSTALLATION

Professors and learners of data mining techniques who would like to use RapidMiner software as a tool to study and experiment for data analysis can download RapidMiner software at the following this link, https://rapidminer.com/, which will take you to the download page as seen below (Chavan, 2020: Soft, 2020: Vasudevan, 2021).

Figure 1. RapidMiner software download
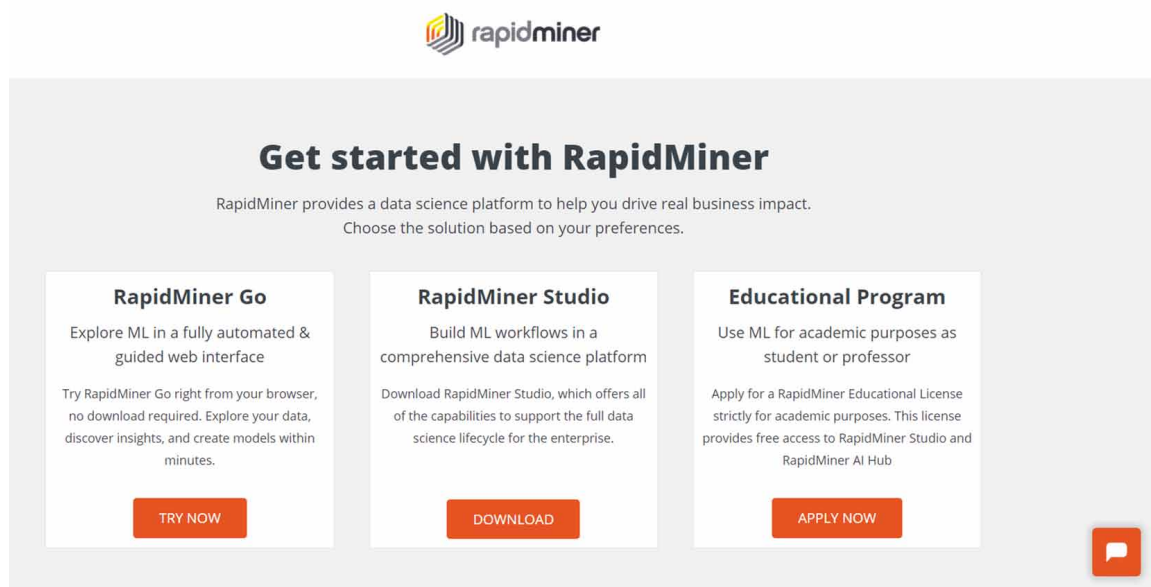
As seen in the figure, select Apply Now at RapidMiner in the form of Education Program, then fill in the details to verify your identity, as seen below.

From figure 3.2, select Educational Purpose, and then fill in your first and last names, educational institution name, the role of the software user consisting of teachers and students, email address and password (Naparat, 2020). Then press the button Register.

As seen, my.RapidMiner.com will send an e-mail requiring the user to confirm the software download request. The user, then, should verify the e-mail.

From the picture, once the email is verified, my.RapidMiner.com will download the software as shown in the picture down here.

From the figure, check the type of operating system inside the computer. In the case of Windows operating system where there are 2 types; 32 Bit and 64 Bit, select the correct download according to the operating system. And in the case of MAC OS or Linux operating system, choose to download according to the operating system. Once the download is completed, you will receive a file for the software installation.

When double clicking on the Installation File, the start screen will appear. Then press the Next > button to go to the next step.

Related Content

Time Series Mining: Background and Related Work
Wynne Hsu, Mong Li Leeand Junmei Wang (2008). *Temporal and Spatio-Temporal Data Mining (pp. 14-43).*
www.irma-international.org/chapter/time-series-mining/30260

An Intelligent Heart Disease Prediction Framework Using Machine Learning and Deep Learning Techniques
Nasser Allheeib, Summrina Kanwaland Sultan Alamri (2023). *International Journal of Data Warehousing and Mining (pp. 1-24).*
www.irma-international.org/article/an-intelligent-heart-disease-prediction-framework-using-machine-learning-and-deep-learning-techniques/333862

Adoption of Dual Iris and Periocular Recognition for Human Identification
R. Deepika, M. R. Prasad, Srinivas Chetanaand T. C. Manjunath (2016). *Research Advances in the Integration of Big Data and Smart Computing (pp. 250-266).*
www.irma-international.org/chapter/adoption-of-dual-iris-and-periocular-recognition-for-human-identification/139406

A Methodology for Building XML Data Warehouses
Laura Irina Rusu, J. Wenny Rahayuand David Taniar (2005). *International Journal of Data Warehousing and Mining (pp. 23-48).*
www.irma-international.org/article/methodology-building-xml-data-warehouses/1750

A Hierarchical Security Model for Multimedia Big Data
Min Chen (2016). *Big Data: Concepts, Methodologies, Tools, and Applications (pp. 441-453).*
www.irma-international.org/chapter/a-hierarchical-security-model-for-multimedia-big-data/150178