Prototyping a Virtual Tutor with Modular Teaching Styles

Andrea Linxen¹, Simone Opel², Stephanie Ebbing³ and Christian Beecks⁴

Abstract: Digitization and artificial intelligence (AI) have entered education in several different ways. While the current development of large language models enables students to access vast knowledge, virtual tutors offer a more tailored approach and better support for time-independent and individualized learning. As several AI systems have become more common in higher education, we developed a concept to implement a virtual tutor with different teaching styles, using the conversational AI platform Rasa. This tutor teaches students basic Exploratory Data Analysis in linear, free, and modular learning phases. In the future, we will conduct comprehensive studies to determine the ability of the virtual tutor to support successful learning.

Keywords: Virtual Tutor, Rasa Framework, NLP, chatbot, learning phases, learning analytics

1 Introduction

Exploratory Data Analysis (EDA) is utilized by data scientists to investigate datasets and discover new insights. Teaching this skill to data science students can be challenging as mastering EDA requires both theoretical understanding and practical experience. While data science programs strive to offer training opportunities, individual tutoring in higher education is restricted by large group sizes and scheduling constraints. Intelligent tutoring systems can help educators to overcome these challenges by offering highly personalized training at convenient times and locations. We developed a concept for a virtual tutor for EDA and implemented a prototype with modular teaching styles with the Rasa framework.

While numerous frameworks are currently available for implementing chatbots, the Rasa framework is particularly suitable for the creation of task-based systems. A task-based virtual tutor is capable of querying user input and processing the information into actions, such as calculating the statistical characteristics of a dataset. Furthermore, the Rasa framework can support the requirements for a virtual tutor, which we identified by analyzing existing tutoring systems [GJ19], [Ho19]: The virtual tutor offers students the opportunity to explore the dataset independently while aiding struggling students. If necessary, the virtual tutor can teach the theory of EDA and provide constructive feedback to students. By monitoring the progress and keeping track of the predefined goals, the virtual tutor can personalize the support for each student.

¹ FernUniversität Hagen, Data Science, Universitätsstr. 11, 58097 Hagen, andrea.linxen@fernuni-hagen.de
² FernUniversität Hagen, Fakultät M+I, Universitätsstr. 11, 58097 Hagen, simone.opel@fernuni-hagen.de
³ FernUniversität Hagen, Data Science, Universitätsstr. 11, 58097 Hagen, stephanie.ebbing@fernuni-hagen.de
⁴ FernUniversität Hagen, Data Science, Universitätsstr. 11, 58097 Hagen, christian.beecks@fernuni-hagen.de

doi: 10.18420/delfi2023-56
2 The Virtual Tutor

The prototype of the virtual tutor guides students through an exploration of an open-access dataset containing sales figures from 45 Walmart stores. The virtual tutor is targeted towards students with an intermediate level of competency in Python. Students analyze the fluctuation of weekly sales figures and develop a hypothesis on how the trend correlates to other variables such as holidays or unemployment numbers. The exploration is performed in three steps: data discovery, hypothesis formulation and hypothesis testing. Corresponding to the three steps of the data analysis, three separate learning phases with three different teaching styles of the virtual tutor are provided. During the first learning phase, students familiarize themselves with the provided data and develop a basic understanding of the contained information. The virtual tutor guides them through the data discovery on a linear path with five predefined steps and offers appropriate support using an integrated back-end dictionary. In the second learning phase, students explore the data in an iterative approach: they examine the connection between the weekly sales figures and other variables to formulate a hypothesis. The virtual tutor adjusts to a new teaching style to allow students more independence. Instead of structured guidance, the tutor offers hints if a corresponding intent is detected. The second phase ends with a short quiz, to verify the hypothesis of the student. Students test their hypotheses in the final learning phase by gaining new insights from the dataset. The virtual tutor focuses on interacting as naturally as possible and offering feedback on the students' progress. The tutor observes the intent of the students' inquiries about the dataset and selects appropriate actions such as feedback. When students indicate the conclusion of their research, the virtual tutor verifies the completion of each goal and either comments on other possible tasks or initiates the finishing.

We examined the viability of the tutor in a user survey. The tutor received positive feedback regarding the different teaching styles, especially the invitation to explore the dataset freely was well received. One user indicated that “[...] the reaction of the tutor depends a lot on asking a question, while it did struggle with requests not formulated as a question.” This problem will be resolved in the next version by adding more detailed intents. The virtual tutor is nevertheless the first step towards a powerful tool that supports students' self-directed learning. The prototype was able to meet the proposed requirements for the virtual tutor. The findings from the usability survey indicate that the task-based system of the Rasa framework is capable of hosting a virtual assistant to assist independent learning. In the future, we plan to conduct studies to examine learning success for data science education in a university setting. Furthermore, we will conduct comprehensive studies to determine the ability of the virtual tutor to support successful learning. In addition, we will continue to develop the virtual tutor and integrate it into a blended learning environment.

Bibliography
