

Title:

New Cytomine modules for user behavior analytics in digital pathology

Year:

2018

Author:

Laurent Vanhee – Master in Computer Sciences with a focus on Intelligent Systems

Background

In-depth behavioral analysis of users of digital pathology software can provide useful insights in education and diagnostic settings. In education, it can help to better understand and improve how students are learning. In diagnostic, it might help to design more efficient recognition algorithms based on experienced pathologist's browsing patterns. This project focused more on students that are learning. Over the past year, 395 students participated in this study. These students followed the course "General histology and alternative experimentation methods that do not use animals" (HISL054) at the University of Liege. For this course, students used the Cytomine web-application to study cellular images. At the end of the semester, they participated in the exams.

Methods

The Cytomine web-based software (<http://www.cytomine.be/>) for user behavior analytics was extended by the developers. The new back-end modules store online user actions (such as user positions in WSI, user clicks on reference annotations, ...) in NoSQL databases and new RESTful web services allow to export them. New analysis modules built in this project automatically generate heat maps, scan paths, timelines, and more than 2000 behavioral features for sets of users and series of whole-slide images. The tool was combined with tree-based machine learning algorithms (from scikit-learn) to identify most prominent features using the exam results as a basis for analysis. Figure 1 gives an overview of the tool built.

Results

The image heat maps and scan paths provide good means of analyzing user behavior for a single user and image pair. Unfortunately, it is not enough for analyzing behavior as a whole. An ExtraTree regressor machine learning algorithm was implemented. The algorithm was paired with the behavioral features and the global exam results for the entire set of students. Multiple plots were generated to interpret the results and which features were the most impactful. To test the validity of this model, a leave-one-out cross validation was executed. With a mean absolute error of 1.94, the results were acceptable considering the small sample size. Figure 2 shows the predicted grades compared to the original grades. With these results, confirming features and images that stand out is less risky.

Conclusions

This tool provides the means to analyze behavioral analysis of users of digital pathology software. With the many visual ways to analyze the data and the results of the tree-based learning algorithms, it is possible to make observations and assumptions on the user behavior. This tool designed to be used with Cytomine in the future can be used for different studies outside of pathology.

Figures and tables:

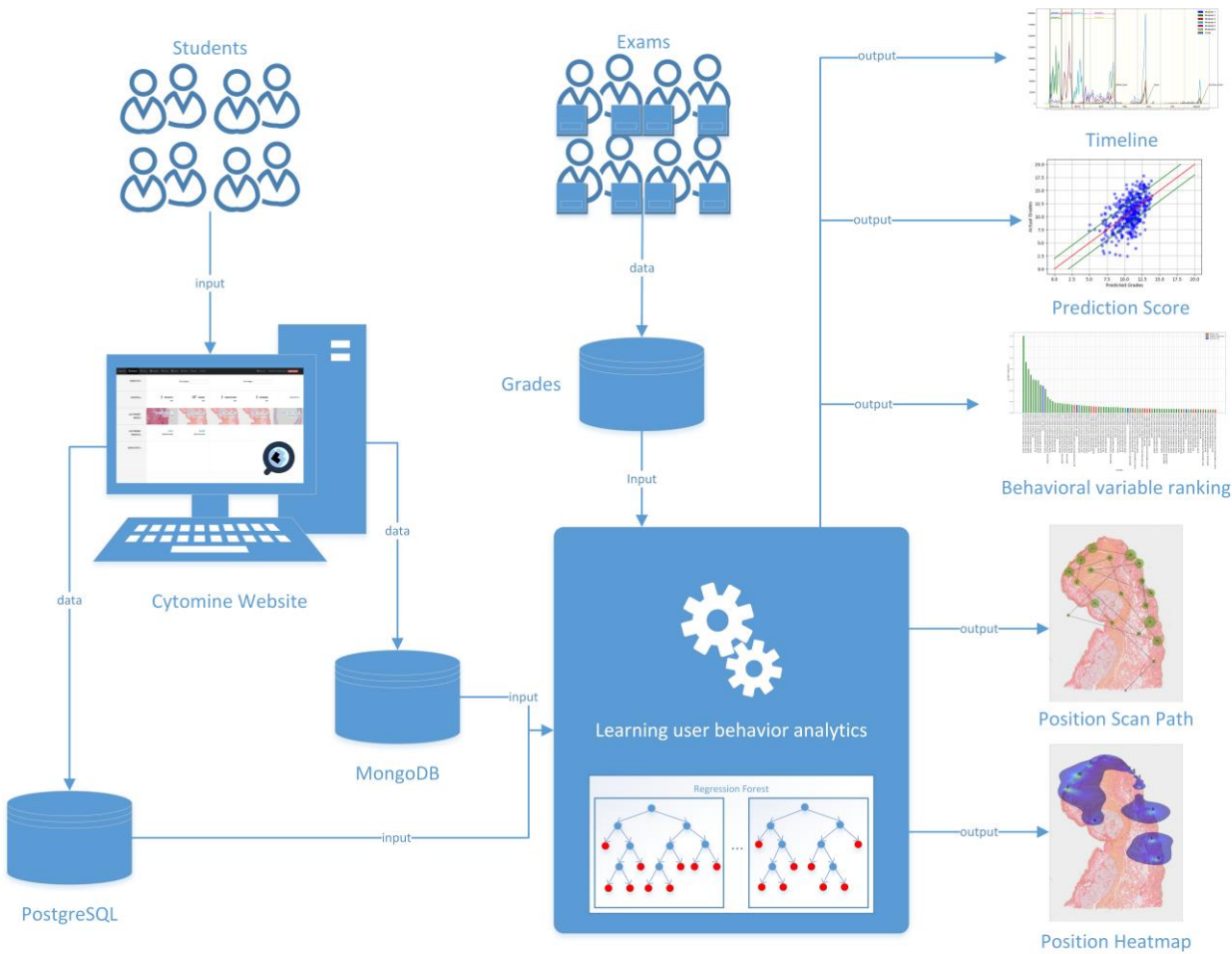


Figure 1: Overview of the Cytomine modules for learning user behavior analytics.

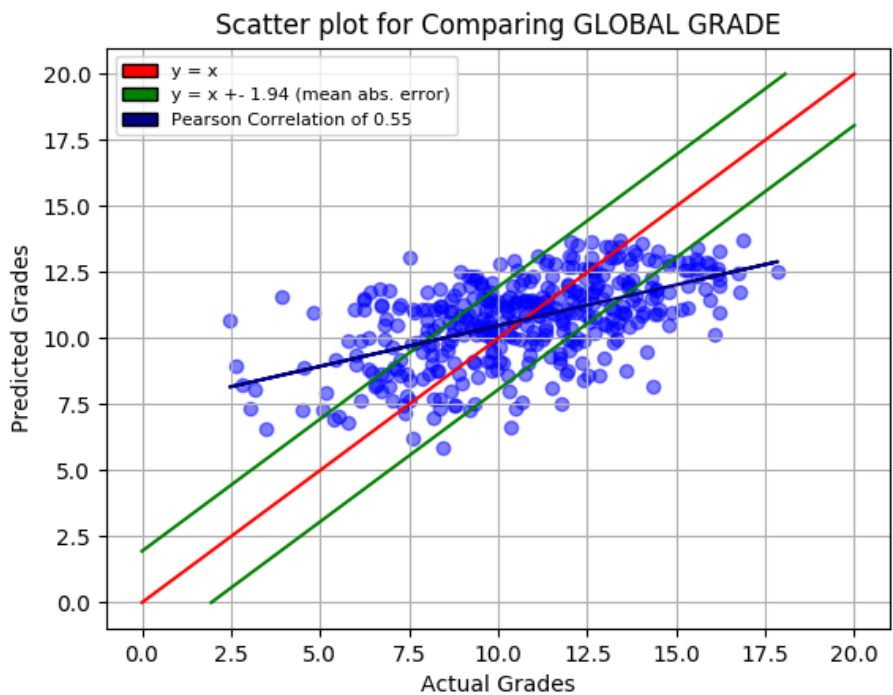


Figure 2: Results of the leave-one-out-procedure for the prediction of global grades.