Personal Visualization + Learning Process

Yannick Lang
Matriculation Number: 1995498
M. Sc. Applied Computer Science

Master Seminar Information Visualization
Visualizing Personal Data

University of Bamberg
Summer Term 2024

1 Introduction & Background

The goal of this paper is to give an overview of visualizations that can be integrated into a student dashboard to improve student success. The focus will be on a student's entire academic career as opposed to single courses. Student success is understood to be the ability of a given student to graduate from their respective degree program in time and with good grades.

There are several phases in which students could be assisted by such visualizations:

1. degree program planning
2. course selection
3. suggestion of interventions in case of problems
4. feedback on past courses

Of course, those phases do not need to be in this order and might repeat or overlap.

The next section will give a brief overview of the types and sources of data that are available, followed by examples of visualizations that can be used to support students in their academic careers. In the end, potential challenges and benefits will be discussed.

2 Available Data

One key question of course is which kinds of data are available and from which sources. In the following, there will be a brief overview of the data sources that are generally available, with a focus on the local situation at the University of Bamberg.

In general, there are two kinds of data: direct data and derived data. Direct data is data that can be collected directly and usually is already available, such as the grades of a student, specifications of degree programs, results of mid-term evaluations, or the interaction of a given student with published course material.

Derived data is data that can be generated from direct data, such as the estimated difficulty of exams, pass rates, the estimated trajectory of a student (if and when they will finish their degree), or the estimated workload of a student. Methods for obtaining such data are considered out of the scope for this paper.

When implementing such a dashboard, one of the key issues is the availability of data. Which kinds of data are available and from what sources is highly dependent on the institution. In the case of the University of Bamberg, several sources of data would need to be combined to create a comprehensive dashboard. The following list is not exhaustive, but gives an overview of the most important sources of data:

FlexNow is where the exam results are stored. This is a highly important source of data, as it contains the grades of all students for all exams. This data can be used to calculate the average grade of a student, the pass rate of a given exam, or the estimated difficulty of an exam.

VirtualCampus is the learning management system of the University of Bamberg. It could be used to gather information about the course material, and the interaction of a student with the course material. This data can be used to estimate the workload of a student and the engagement of a student with the course material.

On **UniVis**, information about all courses is stored. Together with the degree program specification, this data could be used to suggest courses, however, information about the degree programs is not easily accessible, as will be discussed in section 4.

3 Visualization Examples

In this section, some examples of visualization from the literature will be discussed. The visualizations are grouped by their intended use case for the dashboard.

3.1 Success Prediction

The first group of visualizations is concerned with predicting student success.

One quite intuitive design element here is a traffic light scheme, with green indicating a high likelihood of success, yellow indicating a moderate likelihood,

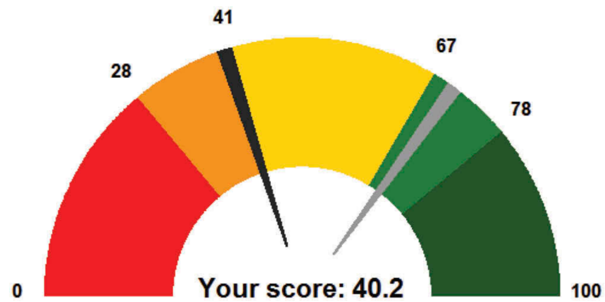

Figure 1: Success score gauge plot, from [5]

and red indicating a low likelihood. This can be seen in the success score gauge plot in Figure 1.

In this plot, there are multiple zones, indicating different risk levels. This plot focuses on derived data and can be used to highlight the likelihood of failure for a given course or the degree program as a whole. Additionally, a second indicator can be added to this plot, to highlight the effect of an intervention or a change in study behavior.

Success prediction can also be helpful when planning which exams to take. In Figure 2, a student can select which exams they intend to take, and the system will predict the likelihood of passing all selected exams. This can help students to focus on fewer exams instead of taking all exams at once.

3.2 Trajectory

Another group of visualizations is concerned with the trajectory of a student. These can be used to show the path, a student has taken so far and is likely to take in the future. This can be helpful to identify students who are at risk of dropping out or to suggest interventions to students who are likely to take longer than expected to finish their degree.

In Figure 3, a selection of typical trajectories is shown. These have been identified by the Authors of [3] and experienced educators. The trajectories might be used to identify students at risk of dropping out, by comparing their actual grade history to the typical trajectories. There remains the risk, however, that the pattern might become apparent only after the optimal intervention time has passed already.

Figure 4 shows multiple trajectories for the same student. The top left plot shows the trajectory over all courses, while the other panels plot achievements for different categories of courses. This can be used to get more nuanced insights into the performance of a student, who might be doing well in some categories, but poorly in others. The insights could be used to inform decisions about which actions to take: for example, non-mandatory courses could be selected from the areas in which the student performs better, the student could invest more time

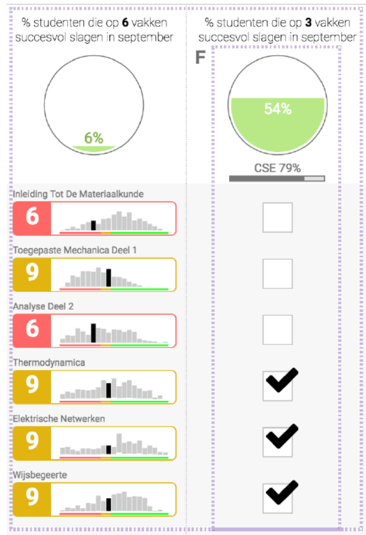

Figure 2: Success prediction for selected exams, from [2]

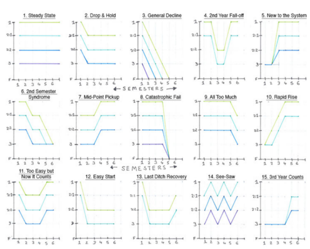

Figure 3: Selection of typical trajectories, from [3]

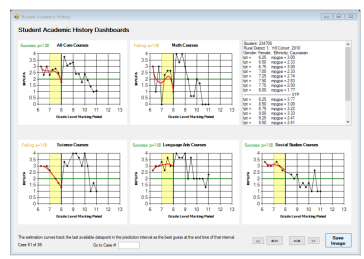

Figure 4: Trajectory of a student for different course bundles, from [7]

into the courses in which they perform worse or a possible intervention could target the categories where the student is struggling.

In Figure 12, estimates for the completion of the study program are shown. Compared to students with a similar profile, this shows the proportion of students that finished their degree program in three, four or five years and the proportion of students that dropped out. This can be useful in the planning phases, as the differences can be compared when selecting different courses for example. Additionally, the influence of passing or failing a given exam can be visualized in such a way.

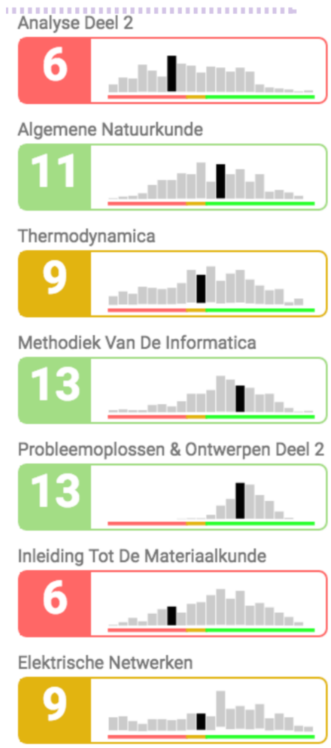

Figure 5: Distribution of grades for multiple exams, from [2]

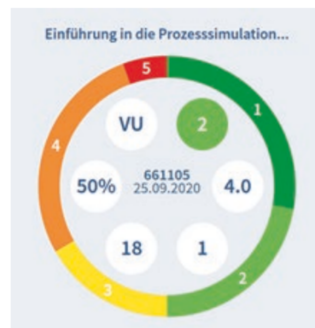

Figure 6: Distribution of grades for a given course, from [8]

3.3 Comparison with cohorts

The Comparison with other students is a particularly interesting area. However, as will be discussed in Section 4, some things need to be considered when comparing students with each other.

There are multiple dimensions, in which a comparison to other students could be helpful. The first one is the exam-level comparison, where the performance of a given student is compared to other students who took the same exam. This can also give an indication of the difficulty of the exam. Figure 5 shows a histogram of grades for a given exam. Additionally, grades that constitute a pass or a fail are highlighted and the grade actually achieved by the student is shown. Figure 6 displays similar data differently: the distribution of final grades for a course is displayed as a donut chart.

Comparisons with other students can also play a role during the planning phase. Figure 7 shows the amount of ECTS credits a given student is currently on track to achieve and the average amount of ECTS credits achieved by other students. Additionally, thresholds are highlighted, that need to be cleared in order to meet certain requirements. This could also be adapted to the specific

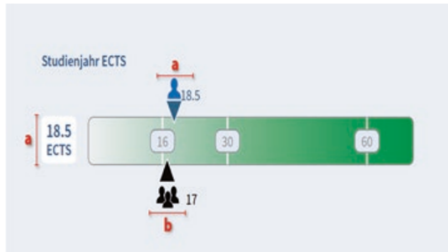

Figure 7: Comparison of ECTS credits achieved by a student and other students, from [8]

Figure 8: Trajectory of a student and comparison with other students, from [2]

requirements of a given degree program.

Of course, some visualizations can belong to multiple groups. For example, Figure 8 shows both the trajectory of a student and the performance compared to other students. The trajectory is shown as a line plot, connecting the grades of the students over time. The comparison with peers is achieved through vertical histograms, akin to violin plots. This allows the user to quickly see how they are doing compared to other students.

So far, all visualizations were concerned with course-level data such as grades or ECTS. However, visualizations can also be used to compare the engagement of students with the course material. Figure 9 shows a radar plot with different dimensions of interaction. Other than the line for a given student, other thresholds could be shown here. In this example, the average interaction and the student with the highest interaction scores are highlighted. In theory, this could also take some thresholds into account. Some courses have special restrictions on who is allowed to take the final exam, for instance, it might be required to complete 50% of assignments given out. More generally, such plots can be used to identify students who are not engaging with the course material enough and thus might be at risk of failing the exam. Figure 9 also includes recommendations for actions a student could take to increase performance, such as submitting homework on time or retaking a test, in which the student scored lower than usual.

Actual methods for predicting student success are out of scope for this paper, but [10] shows different predictive models based on the interaction of students with the course material, and suggests good predictive power after about 25% of the duration of the course, which should leave time for students to adapt their behavior.

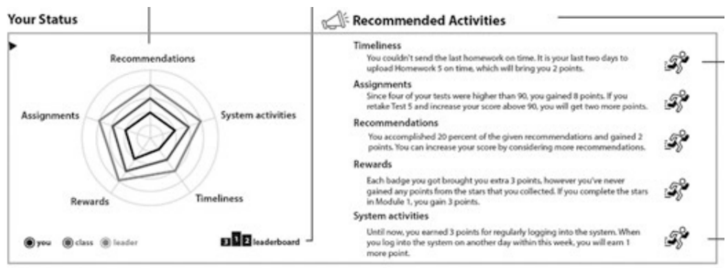

Figure 9: Radar chart with engagement metrics on the left-hand side, and recommended actions to increase performance on the right, from [1]

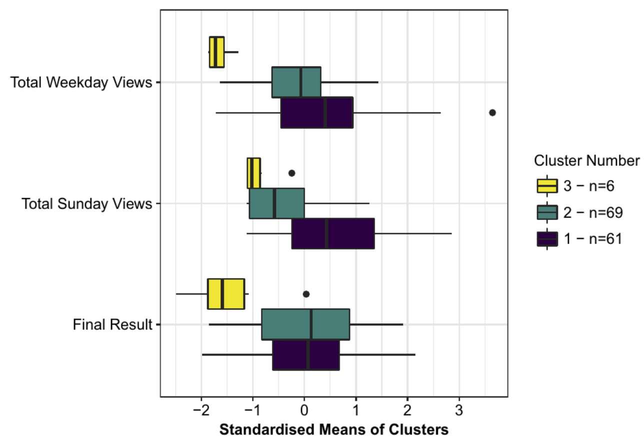

Figure 10: Clusters of students based on their interaction with the course material, from [6]

One thing to keep in mind when comparing students is choosing the correct comparison group. For any given course, often there are many students who are enrolled in the online course but do not participate or decide not to take the exam very early on. Comparing students who actually want to take the exam with those passive students would bias the results. This could lead to a situation where students feel too safe because they are doing more than the average student, but most students doing less are not even planning to take the exam. Figure 10 shows an analysis of different clusters of students, based on their interaction with the course material. This shows that there are distinct groups of students, both with regard to their interaction with course material and their performance, and that interaction with the course material can be used as a predictor of performance, however, the relationship is not necessarily linear. Such classification should thus be performed before comparing students with each other, to help students make better decisions.

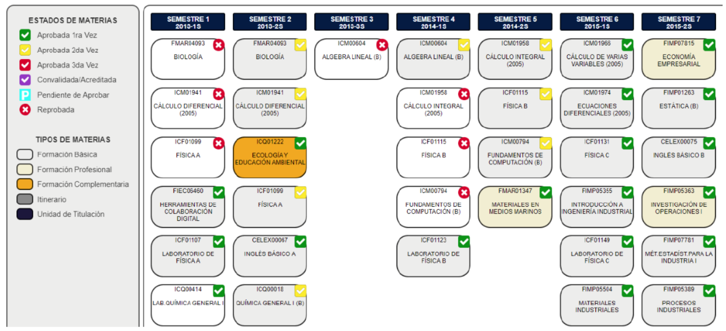

Figure 11: Course selection with indicators showing the year that they are appropriate for, from [4]

3.4 Planning

There are multiple phases in a degree program what involve planning. Students might plan their entire degree program upfront, or they might choose which courses to take in the next semester. Sometimes, students also might have to reconsider if they will be able to pass all exams for the current semester or if they should drop some courses, instead focusing on achieving better grades in fewer courses.

It would be interesting to see, how visualizations could be used to select courses for an upcoming semester, but interesting examples of this are scarce in the literature. Most visualizations do little more than list available courses, with colors to indicate if a course is mandatory or not [8].

Visualizations intended for planning the entire degree program duration might additionally take into account prerequisites for courses. The visualization that comes closest to this is shown in Figure 11. Although prerequisites are not explicitly shown, colors indicate which courses are intended for which year of the study program. Such visualizations can also take into account the number of ECTS credits that need to be achieved in a given semester, an approximated workload for a given course (which in theory should match the ECTS credits, but in practice there usually are quite large differences between courses with the same credits) and the availability of courses. Then, students can plan their degree program in a way that balances their workload while still achieving the necessary credits in a given time frame.

Figure 13 shows, how such a focus on the credits could look like, with the entire amount of credits that need to be gained (180ECTS for a Bachelor's degree) being displayed as a progress bar up top and multiple sliders for the individual semesters below. Additionally, visualizations such as in Figure 5 and 12 can be useful to inform the planning process.

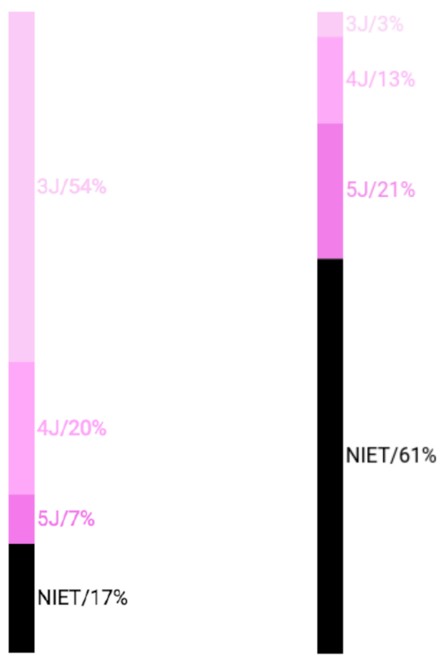

Figure 12: Estimates for completion of the study program, from [2]

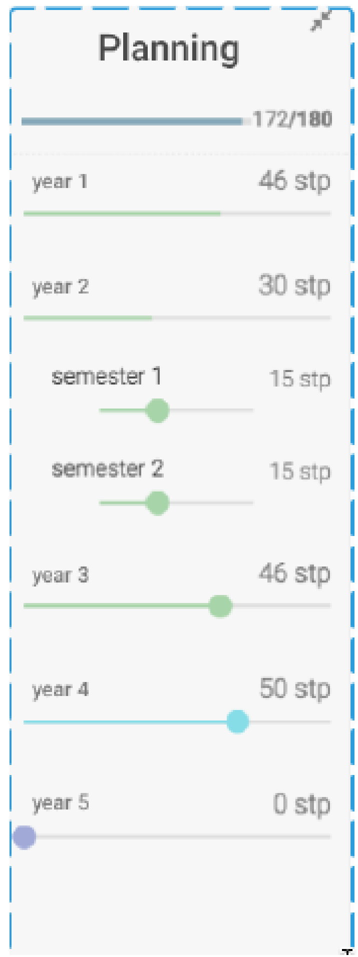

Figure 13: Allocation of credits to the semesters, from [9]

4 Challenges

Of course, several challenges need to be addressed when implementing such a dashboard. The most important challenges are: the availability of data, privacy concerns, and effectiveness.

4.1 Availability of Data

When implementing such a dashboard, one of the key issues is the availability of data. Which kinds of data are available and from what sources is highly dependent on the institution. As discussed in Section 2, in the case of the University of Bamberg, several sources of data would need to be combined to create a comprehensive dashboard, and it is not clear that all data is already readily available.

Especially for the planning aspects, it is important to have data about all courses and degree programs available. Usually, degree program specifications are handed out as PDFs, which are not easily machine-readable. This data would need to be converted into a machine-readable format to be used in a dashboard. The system would also need to handle changes to degree programs, which might apply to every student (because some courses are no longer offered) or only to new students (because the degree program has been adapted). Some universities also offer degree programs in cooperation with other universities, which would make the data gathering even more complex [8].

4.2 Privacy Concerns

There are multiple ways in which privacy and data protection concerns could arise. For example, if the dashboard is used to suggest interventions in case of problems, it is important to ensure that the data is only accessible to the student and the responsible staff members. In general, it is important to ensure that the data is only accessible to the people who need it and that the data is not misused. Students should be allowed to decide who can see their data and should be able to revoke access at any time. This is especially important when the data is used to suggest interventions, as this could be seen as an intrusion into the private life of the student.

Especially visualizations in which the student is compared to others are a sensitive topic. The data shown to the student must not allow for the identification of other students. For small courses, this might make it impossible to show certain visualizations, as another student would be easily identifiable. Therefore, comparative visualizations need to strike a balance between showing useful information and protecting the privacy of the students. This is especially important when students are compared to cohorts instead of the entire student body, which, as discussed previously, might improve the information that can be gained from the visualization but will also decreas the anonymity set.

4.3 Effectiveness

The most important challenge is to ensure that the visualizations are actually effective. This means that the dashboard should help students to succeed in their academic career. One very important thing to consider is motivation of students. If the visualizations are too negative, students might be tempted to give up. For example, if a student is shown that they are behind the average of their cohort, with a small chance of passing the exam, they might become demotivated and drop the class instead of making an effort and working harder. On the other hand, if they know that they are ahead of the average, they might become complacent and work less, thus actually decreasing their chance of success. Therefore, it is very important to conduct field tests to ensure that the visualizations are not demotivating students.

If the dashboard is used to suggest interventions or to urge students to work more, it is important to ensure that such feedback comes at the right time. Early in the semester, students will still have enough time to change their work habits and catch up, while later in the semester, it might be too late to make a difference. On the other hand, there might not be enough data that soon to verify, whether any changes are needed at all. Therefore, it is important to ensure that the feedback is timely and actionable.

5 Benefits

The benefits of a student dashboard with such visualizations can be manifold. The most important benefits are: a good overview for students, easier planning, early detection of problems, personalized feedback, and increased motivation.

When executed well, such a dashboard could reduce problems from the very beginning of the academic career, by allowing students to plan their degree program more effectively. If and when problems arise, students could be supported by personalized feedback and suggestions for interventions, hopefully motivating them to participate and study more, thereby increasing their chances of completing all courses. If a student does fail an exam, they could be assisted in planning their next steps. Depending on the type of course, this might mean choosing when to retake the exam or choosing a different course that is more suitable for them.

There are benefits of such a dashboard that not only apply to students but also to the entire university. Depending on how exactly the dashboard is implemented and data privacy concerns are mitigated, instructors and administrators could also benefit from the insights gained from the data as well as from higher graduation rates.

Another potential benefit that has not been discussed so far is the social aspect of studying. While this would certainly have even more privacy implications that would need to be addressed, it could be beneficial for students to not only see, how they perform compared to peers but also to see how they could help each other. For example, if a student is struggling with a course, they could be shown other students who have already passed the course and could be contacted for help. In the pre-semester planning phase, it could also be interesting to take into account what courses other students are planning to take, to make it easier to find study groups, especially if those students have a similar profile or there have already been some common classes. However, I am not aware of any research that has been done in this area, so this is purely speculative.

6 Conclusion

This paper has given an overview of visualizations that can be integrated into a student dashboard to improve student success. The focus was on visualizations that can help students plan their entire academic career and adapt to changes if necessary. The visualizations can be broken down into four categories: degree program planning and course selection, success prediction, trajectory, and comparisons with other students. The paper also discussed the availability of data, potential challenges, and benefits of such a dashboard. Potential for future work lies in the implementation of such a dashboard, especially in the evaluation of its effectiveness with regard to increasing the motivation of students to perform as well as possible. There could also be value in creating new visual methods for selecting courses, taking into account the workload and prerequisites of the courses.

References

[1] Fatma Bayrak, Pınar Nuhoğlu Kibar, and Selay Arkün Kocadere. "Powerful Student-Facing Dashboard Design Through Effective Feedback, Visualization, and Gamification". In: *Advances in Analytics for Learning and Teaching* (2021). URL: `https://api.semanticscholar.org/CorpusID:245270369`.

[2] Sven Charleer et al. "Learning Analytics Dashboards to Support Adviser-Student Dialogue". In: *IEEE Transactions on Learning Technologies* 11 (2018), pp. 389–399. URL: `https://api.semanticscholar.org/CorpusID:52897854`.

[3] Cameron C. Gray, Dave Perkins, and Panagiotis D. Ritsos. "Degree pictures: Visualizing the university student journey". In: *Assessment & Evaluation in Higher Education* 45 (2019), pp. 568–578. URL: `https://api.semanticscholar.org/CorpusID:214350174`.

[4] Julio Guerra et al. "Adaptation and evaluation of a learning analytics dashboard to improve academic support at three Latin American universities". In: *Br. J. Educ. Technol.* 51 (2020), pp. 973–1001. URL: `https://api.semanticscholar.org/CorpusID:220530168`.

[5] Lingjun He et al. "Predictive Analytics Machinery for STEM Student Success Studies". In: *Applied Artificial Intelligence* 32 (2018), pp. 361–387. URL: https://api.semanticscholar.org/CorpusID:52298674.

[6] Emma Howard, Maria Meehan, and Andrew C. Parnell. "Contrasting prediction methods for early warning systems at undergraduate level". In: *Internet High. Educ.* 37 (2016), pp. 66–75. URL: https://api.semanticscholar.org/CorpusID:65483415.

[7] Ph.D Warren E. Lacefield and Ph.D E. Brooks Applegate. "Data Visualization in Public Education: Longitudinal Student-, Intervention-, School-, and District-Level Performance Modeling." In: 2018. URL: https://api.semanticscholar.org/CorpusID:149659364.

[8] Philipp Leitner et al. "Visualization of Learning for Students: A Dashboard for Study Progress – Development, Design Details, Implementation, and User Feedback". In: *Advances in Analytics for Learning and Teaching* (2021). URL: https://api.semanticscholar.org/CorpusID:245286522.

[9] Martijn Millecamp et al. "A qualitative evaluation of a learning dashboard to support advisor-student dialogues". In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (2018). URL: https://api.semanticscholar.org/CorpusID:3695776.

[10] Pedro Manuel Moreno-Marcos et al. "Temporal analysis for dropout prediction using self-regulated learning strategies in self-paced MOOCs". In: *Comput. Educ.* 145 (2020). URL: https://api.semanticscholar.org/CorpusID:208093100.