

# Direct automated feedback for student submissions based on LLMs

MAXIMILIAN SÖLCH, Technical University of Munich, Germany

FELIX DIETRICH, Technical University of Munich, Germany

STEPHAN KRUSCHE, Technical University of Munich, Germany

A clear and well-documented  $\text{\LaTeX}$  document is presented as an article formatted for publication by ACM in a conference proceedings or journal publication. Based on the “acmart” document class, this article presents and explains many of the common variations, as well as many of the formatting elements an author may use in the preparation of the documentation of their work.

CCS Concepts: • **Social and professional topics** → **Student assessment**; • **Applied computing** → **Education**.

Additional Key Words and Phrases: Do, Not, Us, This, Code, Put, the, Correct, Terms, for, Your, Paper

## ACM Reference Format:

Maximilian Sölch, Felix Dietrich, and Stephan Krusche. 2024. Direct automated feedback for student submissions based on LLMs. In . ACM, New York, NY, USA, 3 pages. <https://doi.org/XXXXXXX.XXXXXXX>

## 1 INTRODUCTION

describe the problem of providing feedback to students in programming courses time consuming, not scalable, not always available hinders learning progress and motivation  
outline of the paper at the end

## 2 RELATED WORK

[1]

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

### 3 APPROACH: DIRECT AUTOMATED FEEDBACK DELIVERY (DAFEED)

DAFeeD employs large language models to deliver automated feedback on student submissions, designed to complement traditional teaching methods and provide additional support. Figure 1 illustrates the continuous feedback workflow that DAFeeD facilitates, enabling students to receive feedback at any time, thereby eliminating the need to wait for responses from human tutors or course professors.

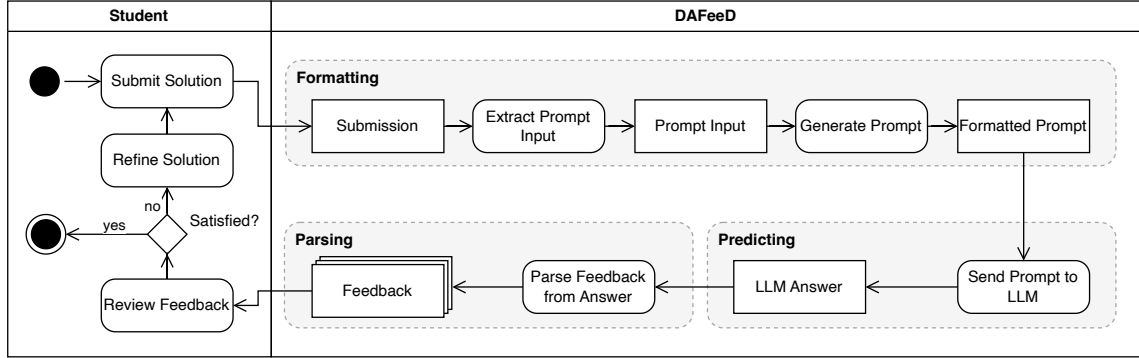


Fig. 1. Workflow of direct automated feedback delivery for students' submissions (UML Activity Diagram)

DAFeeD can provide feedback on various aspects such as the correctness of the code, the quality of the code, and the performance of the student. Once the student submits their solution, DAFeeD initiates a three-stage process to generate natural language feedback.

The first stage, called *Formatting*, takes the student's submission and extracts the submission content, problem statement including learning objectives, and any possible grading instructions the instructor defines. This extracted information represents the prompt input. During the prompt generation step, a predefined prompt template is filled with the prompt input data, resulting in a formatted prompt.

In the second stage, called *Predicting*, the formatted prompt is sent to a large language model, which generates a response that includes detailed feedback for the student.

The final stage, *Parsing*, takes the LLM response, which comes in the JSON format, and parses feedback items from it. In addition to the feedback text, the feedback object also contains reference information indicating the part of the submission it pertains to. For programming exercises, this includes the file name and line number of the relevant code snippet to which the feedback refers. For text exercises, the reference information includes only the sentence or word range the feedback refers to.

All of the feedback is then returned to the student for review. If the student is satisfied with the feedback, the process concludes. Otherwise, the student can refine their solution and resubmit it, initiating the DAFeeD process anew.

This iterative process is designed to motivate students to continuously learn and experiment with their solutions, resulting in improved performance.

## 4 REFERENCE IMPLEMENTATION: ATHENA

### 4.1 Prompts

### 4.2 Feedback Generation

### 4.3 Architecture

Modular approach to support multiple exercise types and extend it in the future

The assessment module manager receives all requests, checks for authorization, and then forwards them to the responsible modules.

Independent of a specific learning management system, provides a REST API documented with the OpenAPI standard

Currently connected to OpenAI and Azure OpenAI Can be simply modified to use open-source models like Llama or Mistral either self-hosted or in the cloud.

As the reference implementation is used in conjunction with the Artemis learning management system for actual university courses, the system is designed to be scalable and reliable (performance, maintainability, usability) Deployed in a kubernetes cluster to ensure scalability and reliability, loadbalancing, etc.

## 5 EVALUATION

### 5.1 Research Questions

### 5.2 Study Design

### 5.3 Results

### 5.4 Limitations

### 5.5 Discussion

## 6 CONCLUSION

Future work improve visualization of feedback, improve feedback quality implement for more exercise types Use direct automated feedback in a real-world setting by enabling the feature in a real course. This allows us to collect more data to evaluate the impact on student performance and motivation.

## REFERENCES

- [1] Marcelo Guerra Hahn, Silvia Margarita Baldiris Navarro, Luis De La Fuente Valentin, and Daniel Burgos. 2021. A Systematic Review of the Effects of Automatic Scoring and Automatic Feedback in Educational Settings. *IEEE Access* 9 (2021), 108190–108198. <https://doi.org/10.1109/ACCESS.2021.3100890>