



Challenge: PDF Highlight and Find Related Text

Feature requirements for Fullstack Engineer

1. Introduction

Build a small web app that allows a user to view a PDF, highlight selected text, find related text elsewhere in the same PDF, and jump to and visually indicate those related locations. Suggest using Programming Language: JavaScript, TypeScript, Python.

2. Feature Requirements

2.1 Frontend (FE)

PDF Viewing

- Display a PDF document in the browser.
- Support basic navigation (scroll and/or page controls).

Highlight

- Allow the user to select text and create a highlight.
- Show the highlight visually on the PDF.
- Maintain a simple list of created highlights (e.g., in a side panel).

Find Related

- Provide an action for a highlight: “Find related”.
- Display results in a panel/list. Each result must include: page number, a short snippet/preview, and a “Jump” action.

Jump + Visual Indication

- When the user clicks “Jump”, navigate to the matched location in the PDF.
- Visually indicate the matched location (e.g., box/overlay/highlight) so the user can see where it is.

States



- Show a loading state while searching.
- Show an empty state if no related results are found.

2.2 Backend (BE)

Related Text API

- Provide an API endpoint to search for related text within the same PDF.
- The API input must include the highlighted text (query) and the PDF content (directly or indirectly—candidate's choice).
- Return a ranked list of related matches within the PDF.

Response Requirements

- Each returned match must include enough information for the FE to: show a snippet, know the page number, and locate + visually indicate the matched area in the PDF.
- Exact request/response format is up to the candidate, as long as the FE can implement Jump + Visual Indication reliably.

Non-functional

- Reasonable performance for typical PDFs (tens of pages).
- Basic error handling with meaningful error responses.

3. Deliverables

- Frontend + backend source code.
- README with run instructions.
- Brief explanation of:
 - (a) how related results are computed.
 - (b) how the FE locates and renders matched areas.

4. Bonus (Optional)

- Provide a confidence score for each match.
- Support multiple highlights.
- Improve matching quality (e.g., definitions, repeated clauses, or better ranking).