Okay, since I cannot directly generate and provide a file, I will describe the **structure and content** of a PDF that would be ideal for testing your model, along with instructions on how you can easily create one yourself using a word processor.

Description of an Ideal Test PDF (for your model)

Title: "The Future of Artificial Intelligence: A Comprehensive Overview"

Page Count: 5-10 pages (well within the <= 50 page limit)

Content Structure:

- **Page 1:**
    - **Title:** "The Future of Artificial Intelligence: A Comprehensive Overview" (Largest font, centered)
    - **Author:** John Doe
    - **Date:** July 23, 2025
    - **H1:** "1. Introduction to AI" (Large font, bold)
        - Brief introductory paragraph about the scope of the document.
- **Page 2:**
    - **H2:** "1.1. What is Artificial Intelligence?" (Slightly smaller font than H1, bold)
        - Paragraph defining AI and its various sub-fields.
    - **H3:** "1.1.1. Machine Learning" (Smaller font than H2, bold)
        - Paragraph explaining ML, supervised/unsupervised learning.
    - **H3:** "1.1.2. Deep Learning"
        - Paragraph explaining deep neural networks.
- **Page 3:**
    - **H2:** "1.2. Historical Milestones"
        - Paragraph on early AI research.
    - **H3:** "1.2.1. Early Concepts (1950s-1970s)"
        - Discussion of Dartmouth Workshop, AI winter.
    - **H3:** "1.2.2. Expert Systems and Revival (1980s-1990s)"
        - Discussion of rule-based systems.

- **Page 4:**
    - **H1:** "2. Current State of AI"
        - Introductory paragraph to current advancements.
    - **H2:** "2.1. Natural Language Processing (NLP)"
        - Paragraph on breakthroughs in NLP.
    - **H3:** "2.1.1. Large Language Models (LLMs)"
        - Specific details on LLMs like GPT, BERT.
- **Page 5:**
    - **H2:** "2.2. Computer Vision"
        - Paragraph on advancements in image recognition.
    - **H3:** "2.2.1. Image Classification"
        - Details on CNNs.
    - **H3:** "2.2.2. Object Detection"
        - Details on YOLO, R-CNN.
- **Subsequent Pages (optional, up to 10 pages total):**
    - **H1:** "3. Challenges and Ethical Considerations"
    - **H2:** "3.1. Bias and Fairness"
    - **H2:** "3.2. Explainability and Transparency"
    - **H1:** "4. Conclusion"
    - **H2:** "4.1. Future Outlook"

Key Features for Testing:

- **Clear Hierarchy:** Distinct H1, H2, H3 levels.
- **Font Variation:** Headings should generally be larger and/or bolder than body text, and H1 > H2 > H3 in size.
- **Page Breaks:** Ensure headings fall on different pages to test page number extraction accurately.
- **Realistic Content:** Uses actual text, not just placeholders.

- **"Pro Tip" Consideration:** You could intentionally make one H3 heading have a slightly larger font than a typical H2, but still be clearly an H3 by its indentation or context, to test your model's robustness beyond just font size.

How to Create This PDF Yourself (using Google Docs or Microsoft Word)

1. **Open a New Document:** Start a new document in Google Docs or Microsoft Word.
2. **Type the Content:** Enter the title, author, date, and all the headings and paragraphs as described above.
3. **Apply Styles:** This is the most important step!
    - For the main title ("The Future of Artificial Intelligence..."), use the "Title" style.
    - For "1. Introduction to AI", "2. Current State of AI", etc., use the "**Heading 1**" style.
    - For "1.1. What is Artificial Intelligence?", "1.2. Historical Milestones", etc., use the "**Heading 2**" style.
    - For "1.1.1. Machine Learning", "1.1.2. Deep Learning", etc., use the "**Heading 3**" style.
    - For the body paragraphs, use the "Normal text" style.
4. **Add Page Breaks:** Insert manual page breaks (e.g., Ctrl+Enter or Cmd+Enter) to ensure headings fall on different pages as described.
5. **Export/Save as PDF:**
    - **Google Docs:** Go to File > Download > PDF Document (.pdf).
    - **Microsoft Word:** Go to File > Save As > Choose PDF as the format.