

Effects of Multimodal Textbooks on Learning

Ananya Chaudhary, Dr. Mahesh Mohan

Abstract— A good teacher can keep students engaged and help them in achieving their learning objectives. However, not every student has access to a good teacher. We propose the use of multimodal learning to keep students engaged and help them learn better. Multimodal learning aims to improve the engagement of every student through the use of different learning modalities. This in turn, will help a larger number of students achieve the overall learning objectives. It is not meant to replace but rather to complement traditional learning methods.

Index Terms— Multimodal Textbooks, learning methods.

I. INTRODUCTION

A multimodal textbook is essentially a textbook that allows for a display of information in a variety of ways such as: text, pictures, audio, and video. This caters to a wider range of students and learning types. While individually, all of these have come to be very commonly used as methods to study, a multimodal textbook utilizes a Machine Learning Model to convert existing education content so its available in multiple other formats. To limit scope, we have picked video as an alternate format.

II. PROBLEMS WITH THE CURRENT METHOD OF EDUCATION

Traditional methods of education - still most commonly in use at schools - are simplistic. They encompass a teacher either reading out excerpts from textbooks and then proceeding to briefly explain them or simply asking students to read certain passages of the book. Alternatively, a teacher will choose to explain a concept and then dictate notes on the same. However, we believe there are some important things wrong with this combination of note-taking, listening, and reading.

A. Unable to maintain the interest of students:

A textbook is a book of instruction. Its primary aim is not to impart information about a specific subject but to enable one to develop a proper understanding of the subject - it is not meant to be comprehensive. However, there is a gap in the learning facilitated by textbooks. Since they are typically written by academic experts and not communication experts or story tellers, there is often a gap in the presentation of the content. They aim to put together as many facts as possible in an organized manner. Further, there are differences in the use of the resource between teachers and students, rendering it less effective. Hence, it is no surprise that traditional textbooks fail to captivate the interest of students.

B. Excessive dependency on listening and reading skills:

While listening and reading skills are imperative to the learning process, visual aids are used to make the learning experience more real, more active, and more accurate. Further, it has been proven that images and visual information are the most effective way of making sure that information is stored long-term. While our short term memory can process words, images are directly processed by our long term memory. Additionally, considering approximately 65% of the population are visual learners and an audience retains 95% of a message when they watch it in a video compared to 10% when reading, the unbalanced focus on verbal skills is not the most effective method to enable learning.

C. Does not cater to different learning types:

It is a well-established fact that different age groups respond differently to different methods of teaching. 4-5-year-olds enjoy songs, colorful images, and movement as well as exaggerated facial expressions and loud sounds. By the age of 6 - 10, children respond more positively to role play and a slight amount of competition. As kids grow older, they begin to want to see a more apparent connection between what they learn and the real world. Hence, the key is to make lessons interactive and relatable. Asking lots of questions is beneficial. Similarly, every child also has a unique learning style and prior understanding of a topic. However, traditional methods treat all children the same. While this often serves its purpose, it also leaves certain kids bored and others lost in class.

D. Occasional visual aids used are loosely related to the topic at hand:

Oftentimes, to cultivate interest, teachers play basic visual aids such as animations and simulations that are loosely connected to the topic but which mainly serve their purpose - to generate hype and excitement. These are not strictly educational as they have no value aside from this. Therefore, they fail to teach students anything and hence interest in the topic is not generated.

III. HOW MULTIMODAL TEXTBOOKS IMPACT EDUCATION

A. The Advantages

While there are many problems with traditional education, there are also many advantages. Hence, the concept of a multimodal textbook - preserving the traditional method of written content in a textbook supplemented by videos explaining the same content is extremely useful. It creates a balance between visual-spatial and verbal skills that is currently lacking, simultaneously making education more

Ananya Chaudhary, Inventure Academy, Bangalore .
Dr. Mahesh Mohan,

interactive and engaging.

B. Immersive learning

Being able to see and hear something is much more effective than being able to only see or only hear. Videos supplemented by textbook content capture the attention of students in a more effective manner. A concept is first explained in a video format and then the student is able to connect concepts they have just understood with what they are reading in the book.

IV. RELEVANT CONTENT IN VIDEOS

Part of the project was creating a Linear Regression model that would be able to predict how relevant a video was when given a subtopic from a textbook. ML models can be used to match high quality, relevant and interesting videos to the pertinent text from the textbook, possibly creating interest in the subject matter.

V. CATERS TO A MIX OF LEARNING TYPES

Since a video has both visual and auditory outputs, it appeals to visual, auditory, as well as verbal learners. Often videos that have certain animations of people on screen also appeal to kinesthetic learners as certain movements are associated with concepts. Further, a multimodal textbook retains the textual content, allowing students that favor reading to still use the resource. Hence, for students who still find it more useful to learn using a simple textbook, they have that option available to them as well.

A. The Disadvantages

1. Inaccessible in certain parts of the world.

Large parts of the world still does not have computers and/or Internet required for this type of teaching to be feasible. Due to this, it could be argued that such methods only widen the already prevalent digital divide.

2. Cause Distraction

Using internet and videos in specific as a study tool can provide distractions. If the videos chosen are of poor quality, then it is more likely for a student to start exploring other things and hence, not learning anything.

VI. HOW TO CREATE A MULTI-MODAL TEXTBOOK

When we started the process, there were many things necessary for us to decide on. Some of the problems we faced were:

- A. **Picking a digital format for textbooks:** We decided to use the EPUB format. EPUB is the most widely supported vendor-independent XML-based (as opposed to PDF) e-book format.
 - a. EPUB is supported by many e-readers and devices.
 - b. EPUB is supported by the International Digital Publishing Forum (IDPF) - a Standards body.

- c. The Book Industry Study Group (BISG) endorsed EPUB format as the "accepted and preferred format" for Web Standards-based digital content.

B. Breaking down content in a textbook into meaningful sub-concepts:

A textbook is organized in chapters. A chapter is split into sub-topics which are further organized into paragraphs. Each paragraph is centered around a core concept. We chose to work with the same structure and embed/ link a relevant video for each sub-topic.

C. For each paragraph, identifying a relevant video:

Considered 2 main approaches:

- a. **Search:** One of the approaches considered was to enter the search query into Google search to find a relevant video. Questions which came up that needed some research were:
 - i. What is the coverage for educational content on YouTube for this to be meaningful?
 - ii. What should be the search query? And is YouTube search meaningful for this use case?
 - iii. What video lengths are available on YouTube and yet useful for students?
 - iv. **Analysis:** Found the relevant video coverage to be satisfactory. Tried an early approach to use the paragraph heading as the Search Query. Seemed reasonable. In the long run, we may want to use summarization techniques to come up with the Search Query. YouTube search is optimized for maximizing video watch time and hence was off in 50+% of the queries. Found that 6 of 11 Queries in English had ≤ 1 relevant results with reasonable length in top 5 results; 9 of 11 Queries in Hindi had ≤ 1 relevant results with reasonable length in top 5 results.
- b. **Text similarity:** 2nd approach is to convert Video to Voice to Text and then do a text similarity between the text retrieved from the Video with the Text from the paragraph and then pick the most relevant Video based on the match. However, found that the accuracy of current tools that convert a video to Voice to Text is low and they do not work for international languages. Given this, decided to not pursue this approach.

Picked **Search** as the final approach based on the analysis above.

- D. **Search Relevance:** Searching on Youtube (a common first step while searching for explanatory videos) returned many irrelevant videos for various concepts of interest. To improve our accuracy, we developed a Machine Learning model to rerank the search results in order of relevance. Our objective function was to maximize the relevance of videos based on the query and the context (student, language etc). Steps we took were:

- a. Manual creation of dataset. For this dataset, created feature vector programmatically. Kept a subset for testing.
- b. Experimentally determining the most useful regression model. Tried many algorithms; features; combinations etc to reduce mean squared error for training and test data set. Tried using the Random Forest Regressor, Linear Regressor, and RANSAC Regressor. However, we determined that Linear Regression gave us the most accurate results. The signals we used were: Number of Dislikes, Edit Distance between the title of the video and the query, Video length in seconds, Number of Likes, Number of Views, Age of the video in seconds, Language, and Search rank.
- c. Applied this model for finding most relevant videos for topics in the book.

multi-modal textbook. Core content in both was identical. Also created a simple, short questionnaire that tested minute details as well as the overall concepts. The kids were then divided into two groups of 15 each. One group - 'Group 1' - was given the normal textbook while the other - 'Group 2' - was given the multi-modal textbook for the first topic - 'Structure of an Atom'. Both groups were given 30 minutes with their respective resources and then administered the test.

There were two major types of questions in the test:

1. Minute Details: Questions that tested aspects of the chapter that were not necessarily stressed upon but were mentioned. 'Who was the scientist who discovered the Cathode Ray Tube had charged particles?'
2. Overall Concepts: Questions that tested major important concepts that were stressed upon in the content. 'How was the Plum Pudding Model proven to be wrong?' These included questions that were diagram/image-based.

The answers to these questions were recorded and analyzed. Here are the results:

VII. THE EXPERIMENT

An experiment was conducted with a group of 30 sixth graders in Bangalore. The framework of the experiment was simple. Chose two topics - 'Structure of an atom' and 'Separation Techniques'. For each of these topics, created a

Metric 1: The Overall Performance Rating of Online Textbook Students (PROT):

Question Number	Group 1 (Multimodal Textbook)	Group 2 (Traditional Textbook)	Question Type	PROT
1	3/15 = 20%	3/15 = 20%	1	0
2	9/15 = 60%	10/15 = 66.7 %	2	-6.7
3	14/15 = 93.3%	12/15 = 80%	2	13.3
4	15/15 = 100%	14/15 = 93.3	2	6.7
5	8/15 = 53.3%	3/15 = 20%	2 (Diagram based)	33.3
6	1/15 = 6.67%	8/15 = 53.3%	1	-46.63
7	11/15 = 73.3	9/15 = 60%	2	13.3
8	14/15 = 93.3%	10/15 = 66.7%	2 (Diagram-based)	26.6
Total	75/120 = 62.5%	55/120 = 45.8%		39.87

= (33.3 + 26.6) = 59.9, signifying much better results by Group 1 once again.

VIII. CONCLUSIONS

The results of this experiment clearly show that overall, Group 1 students performed better than Group 2 students. For Category 1 of questions (specific), Group 2 students performed significantly better than Group 1 as PROT = (0 - 46.63) = -46.63. For Category 2 of questions, PROT = (-6.7 + 13.3 + 6.7 + 33.3 + 13.3 + 26.6) = 86.5, so Group 1 students did much better. Specific to Diagram Based questions, PROT

These results prove that for conceptual understanding and visual teaching, the video textbook is a very efficient method of learning. However, plain text proves to be more efficient at small specific details such as statistics.

However, given the time constraint of 30 minutes, the groups would only have had sufficient time to either

completely read the chapter or watch all the videos of the chapter, not both. Therefore, if a multimodal textbook was used in its entirety - both text and video - it would be extremely effective. Thus, we propose the introduction of Multimodal Textbooks into curricula at young ages such as Grade 5 to boost engagement and aid students in their learning process.

IX. FUTURE WORK:

- We collected training data set manually, so dataset was small. A bigger dataset can help improve relevance further.
- More work can be done to find more features and use better, different algorithms to tune the quality further.
- In this work, we used Paragraph heading as the search query. In the long term, we may want to use summarization techniques to come up with a better search query.
- As ePub work evolves, would want to use better features to make videos easier to discover and interact with.
- This work can also be extended to support other language text books as well.

REFERENCES

- [1] Swaminathan, Saishruthi. "Linear Regression - Detailed View." Medium, Towards Data Science, 26 Feb. 2018, <https://towardsdatascience.com/linear-regression-detailed-view-ea73175f6e86?gi=52c9996a36ee>.
- [2] "Linear Regression¶." Linear Regression - ML Cheatsheet Documentation, https://ml-cheatsheet.readthedocs.io/en/latest/linear_regression.html.
- [3] "HOME." IDRE Stats, <https://stats.idre.ucla.edu/r/dae/robust-regression/>.
- [4] Bikowski, Dawn, and J. Elliott Casal. "Https://Scholarspace.manoa.hawaii.edu/Bitstream/10125/44584/2/22_01_bikowskicasal.Pdf."
- [5] Unesdoc.unesco.org, <https://unesdoc.unesco.org/ark:/48223/pf0000366994>.
- [6] "Read 'Learning and Understanding: Improving Advanced Study of Mathematics and Science in U.S. High Schools' at NAP.edu." National Academies Press: OpenBook, <https://www.nap.edu/read/10129/chapter/8#121>.
- [7] O'Neill, Robert. "Http://Textbookuse.pbworks.com/f/Why_Use_Textbooks.pdf."
- [8] Roser, Max, et al. "Quality of Education." Our World in Data, 17 July 2013, <https://ourworldindata.org/quality-of-education>.
- [9] Raja, R., and P. C. Nagasubramani. "Impact of Modern Technology in Education." Journal of Applied and Advanced Research, vol. 3, no. S1, Oct. 2018, p. 33., doi:10.21839/jaar.2018.v3is1.165.
- [10] Wikramanayake, Gihan. "Https://Www.researchgate.net/Publication/216361364_Impact_of_Digital_Technology_on_Education."
- [11] Benton, Lucy. "Why Visual Engagement Is Essential to Education?" Surfily, 5 Apr. 2018, <https://www.surfily.com/blog/visual-engagement-education>.
- [12] "5 Advantages and Disadvantages of Traditional Education." Essay Mania, 11 Apr. 2017, <https://www.essaymania.co.uk/blog/5-advantages-and-disadvantages-of-traditional-education>.
- [13] Richtel, Matt. "Technology Changing How Students Learn, Teachers Say." The New York Times, The New York Times, 1 Nov. 2012, <https://www.nytimes.com/2012/11/01/education/technology-is-changing-how-students-learn-teachers-say.html>.
- [14] "EPUB." Wikipedia, Wikimedia Foundation, 19 Sept. 2019, <https://en.wikipedia.org/wiki/EPUB>.
- [15] "Comparison of e-Book Formats." Wikipedia, Wikimedia Foundation, 23 Aug. 2019, https://en.wikipedia.org/wiki/Comparison_of_e-book_formats.
- [16] Jewitt, Carey. "Multimodality and Literacy in School Classrooms." Review of Research in Education, vol. 32, no. 1, Jan. 2008, pp. 241–267., doi:10.3102/0091732x07310586.
- [17] Flores, Cameo. "Multimodality in Writing." 2015, doi:10.1163/9789004297197.
- [18] Marchetti, Lorena, and Peter Cullen. "Https://Www.cjv.muni.cz/Cs/Wp-Content/Uploads/Sites/2/2016/02/Cr-11516-Marchetti.pdf."