



## AI Blueprint for an app like “All For good Education”

|  |  |
|--|--|
| <p><b>Problem Statement:</b> “Peru ranked 64th out of 77 countries in the 2018 PISA test results”. One of the most talked about problems in Peruvian education today is the low reading comprehension among primary and secondary school students. The low scores in international tests, coupled with multiple examples in the daily lives of Peruvians, demonstrate this. Text comprehension is an essential skill for the personal and social development of children, as well as for their future academic and work success, but unfortunately, the acquisition of this skill by Peruvian students leaves much to be desired. Peruvian school students have great difficulty when it comes to their reading skills, the encoding and decoding process is very complicated for them, consequently, their comprehension levels are low. And that is how the app “All For Good Education” was born.</p> |  |
| <p><b>User:</b> Peruvians school students of elementary and high school, whether from public or private schools.</p>   |  |
| <p><b>Functionality:</b> The AI system will evaluate and measure a child’s level of fluency and, depending on their category, provide recommendations on topics that can help children with their literacy skills, as well as guide teachers in providing a more personalized education.</p>   | <p><b>How Input → Output creates functionality:</b><br/>The application takes sound data from audio that children record of them reading to detect different problems. Depending on the problems found, the system will subtract points from the total score and classify the person in a low, medium or normal reading level, however, this will not be displayed, but recommendations will be provided for the improvement of their skills according to the results given.</p> |
| <p><b>Input (conceptual):</b><br/>The audios contain the sound of the readings that the children have performed.</p>   | <p><b>X (must be in your data):</b><br/>Student’s scores based on the identification of the following criteria:</p> <ul style="list-style-type: none"> <li>• Very slow reading</li> <li>• Lack of fluency in more complex words.</li> <li>• Syllabic reading.</li> <li>• Phonological equivalence errors (reading “seso” instead of “queso”)</li> </ul>  |

|  |  |
|--|--|
|  | <ul style="list-style-type: none"> <li>● Omitting words or pieces of words.</li> <li>● Adding words</li> </ul>                         |
| <b>Output (conceptual):</b><br>Recommendations given based on the identification of the level at which they are according to the problems encountered during the reading.  | <b>The Label Y (must be in your data):</b><br>The reading level obtained (low, medium or normal) through the evaluation of the audios. |
| <b>Dataset construction (where will instances come from in your data?)</b><br>This application will work mainly with public and private schools data, so, the data set will be built from the audios that the school children will perform from the given reading. There is a text and the corresponding audio for it. We will train the system so that when it detects any problem during the reading process, for example, syllabic reading and slow reading, etc, it will decrease points to the child's total score, in this way, at the end when they have finished reading, the points obtained will be categorized in a level: Low, medium or normal. For each level, designated topics will be assigned, the only information that will be displayed at the end of the evaluation process of the IA system. This will help teachers to provide more specialized instruction and to level all students.   |  |
| <b>Risks:</b> <ul style="list-style-type: none"> <li>● As there are different types of accents, the input may have problems in identifying the reading efficiency.</li> <li>● If we train the AI with an accurate tone of voice, complications may arise when entering audio with a higher or lower tone of voice.</li> <li>● In the case of the population with reading transients such as alexia, dyslexia or hyperlexia, the system would confuse their level of reading efficiency, since they themselves have problems reading or comprehending what is being read.</li> <li>● If we do not train the system to identify the background noise and the voice itself, it may present some problems in identifying the problems encountered in the reading.</li> </ul>   |  |
| <b>How to Mitigate Risks:</b> <ul style="list-style-type: none"> <li>● When collecting training data, the audios submitted should come from students from different parts to ensure that the application captures a wide variety of accents.</li> <li>● The data must be carefully trained on voice tone diversity in order to ensure that there is no error in the system if the person presents a higher or lower voice tone.</li> <li>● Special attention should be paid to students who obtain the lowest reading level, as some of these cases may be people who suffer from reading disabilities. Therefore, in case their grade falls into the low level, a notification will be sent to the teacher's account, so that the teacher can make a more thorough evaluation of the student.</li> <li>● The system must be trained very well to identify the background sound as compared to the audio of the student's voice in order not to make any mistakes that would affect the level at which the student is assigned and the topics he/she will be reinforcing.</li> </ul> |  |

