AI-Driven Podcasts on Swiss National Votes

Graduate

Moritz Schiesser

Initial Situation: The Swiss democratic system allows citizens to participate directly in decision-making through national votes. The issues put to vote are often complex, requiring significant effort from citizens to make well-informed decisions. While the government provides an official information booklet for each vote, its extensive and detailed nature can be difficult for many to engage with.

Recent advancements in AI, particularly in text generation, make it possible to create podcasts that explain these issues in a more accessible and engaging way. This project aims to develop a system capable of generating such podcasts using state-of-the-art AI technologies.

Given the politically sensitive nature of these topics, the system must be designed with extreme caution. Automated outputs that inform political decision-making must be strictly and transparently derived from the official source material.

Approach: This project develops a system that automatically generates podcasts explaining the viewpoints of both political sides in a conversational format. To ensure factual accuracy, entailment prediction is integrated into the generation process.

Small fragments of the official source material are provided to the Large Language Model (LLM) as context, keeping the generated output focused and relevant. Each segment produced by the model is then checked against its source using entailment prediction: if a contradiction is detected, the segment is discarded and regenerated.

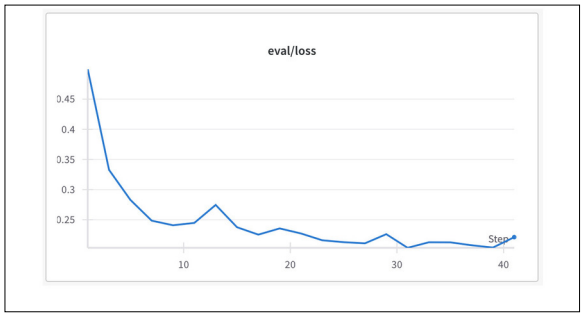
To enable entailment prediction in the German language, an open-weight LLM is fine-tuned for entailment prediction. The entire system is designed with a modular architecture, emphasizing transparency and the use of open-weight models.

Result: The project delivered a functional system that generates AI-based podcasts from the structured information provided in the Swiss voting booklet.

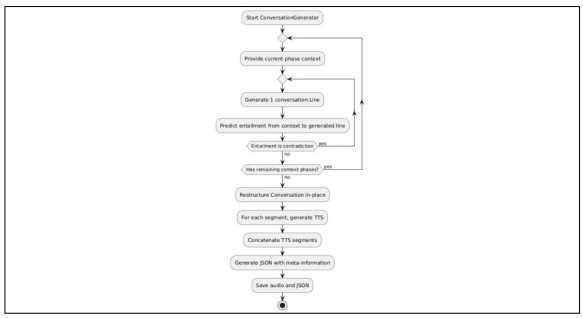
A key achievement is the substantial improvement of an open-source LLM's entailment prediction capabilities through fine-tuning, resulting in a dramatic increase in accuracy. The final product is a web application that allows users to listen to the generated audio while simultaneously viewing the transcript, the original source text, and the entailment prediction for each sentence.

This level of transparency is essential for building trust in the system and its outputs. Furthermore, the modular architecture enables easy updates and enhancements to individual components, ensuring the system remains adaptable to future requirements and advancements in AI technology.

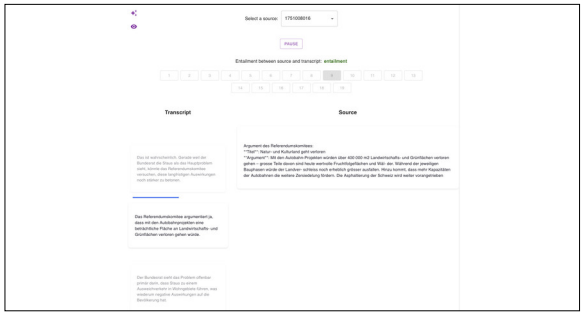
Evaluation loss chart for fine-tuning of entailment LLM
Own presentment



High-level overview of the generation process
Own presentment



The resulting UI showing sections, generated transcript and the source during playback.
Own presentment



Advisor
Prof. Dr. Mitra Purandare

Co-Examiner
Dr. Peter Staar, IBM

Subject Area
Data Science, Computer Science, Software and Systems