

Process (People)

- Scrape articles/transcripts from web
- EDA
- Data processing
- Model training/tuning
- Build App
- Users will take diagnostic test
- User will get recommended articles/videos at their level
- Receive summary, key words, Q&A

Execution (Code)

- Take an article/transcript and apply same preprocessing
- Use the fine-tuned model to make prediction on language level of the text
- Call LLM for article summarization, key words and Q&A generation

State (Source, Data, Models)

- Save articles/transcripts
- Save model weights
- Save LLM-generated samples and features
- Information on pre processing
- Source code repo for collaboration

Solution Architecture

Save the Penguins

