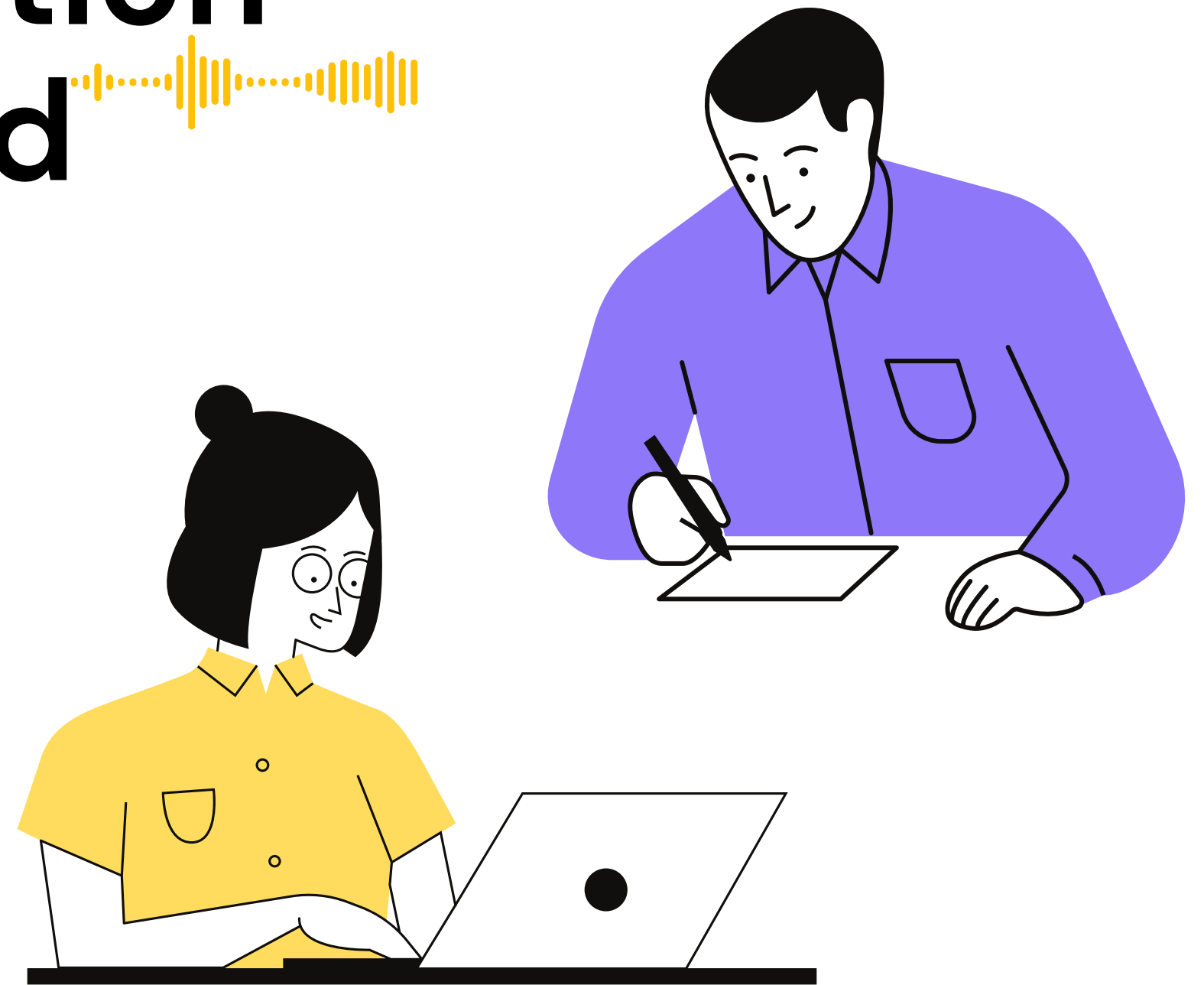


Podcast Summarization Using Wav2Vec2 and BART

For Deeper Understanding and
Efficiency



Meet The Team



KANIA GALIH WIDOWATI
2502047070



KEZIA FOEJIONO
2540131014



M HAIKAL IRAWAN
2502000823

← [BACK TO AGENDA PAGE](#)

Agenda

1

Introduction

2

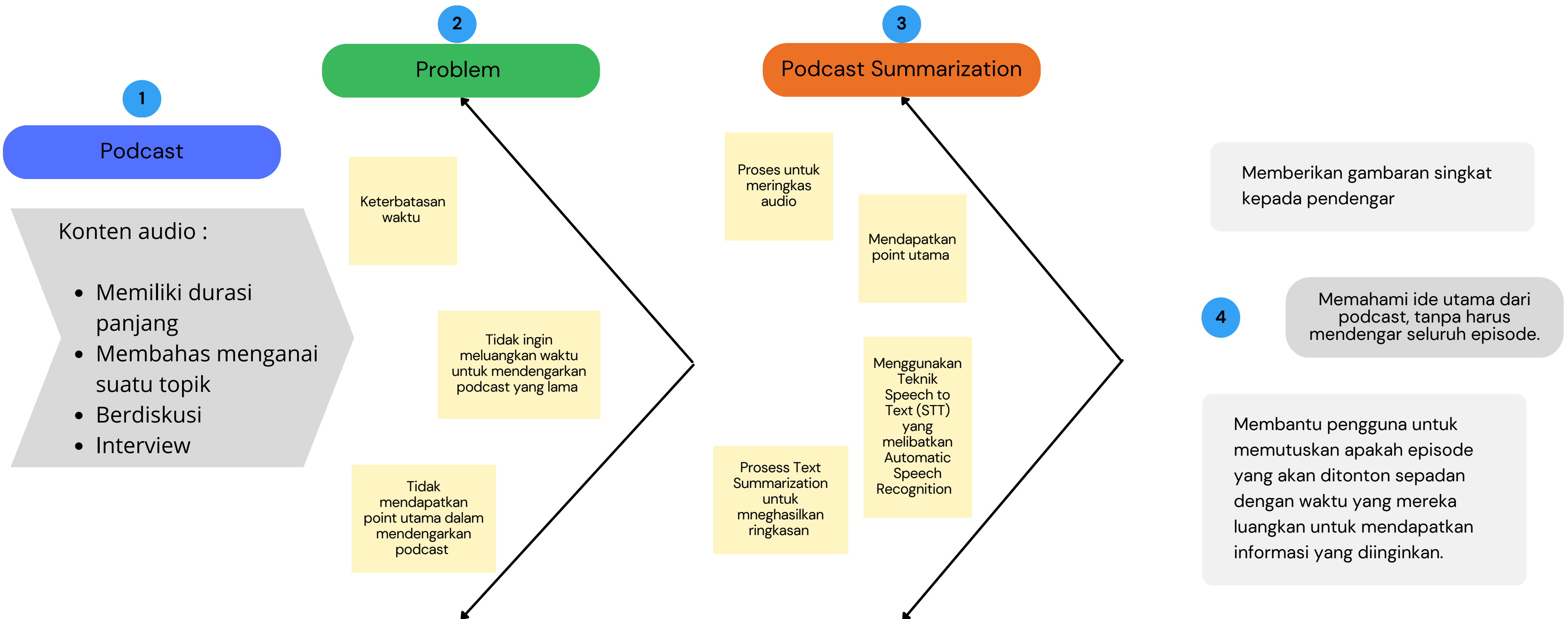
Method & Dataset

3

Workflow

4

Hasil



DATASET

Proses Transkripsi menggunakan Common Voice 6.1 Dataset

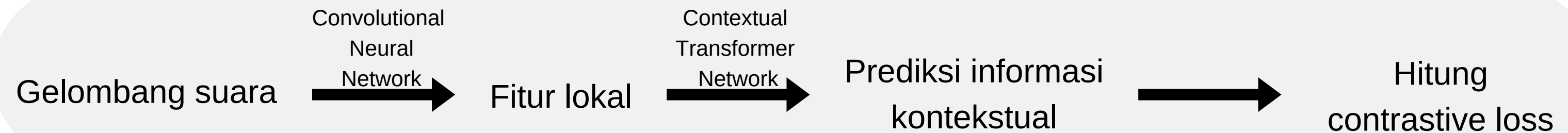
- Disediakan oleh Mozilla
- Terdiri dari 7.335 jam data suara
- Menyediakan 60 bahasa
- Menyertakan metadata demografis (usia, jenis kelamin, aksen)

Proses Summarization menggunakan CNN/Daily Mail Dataset

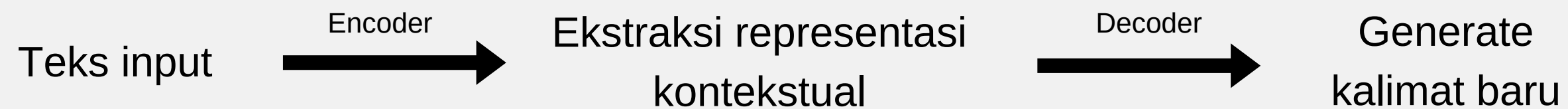
- Menggunakan bahasa Inggris
- Terdiri dari 300.000 artikel berita yang berbeda

METHOD

TRANSCRIPTION menggunakan WAV2VEC2

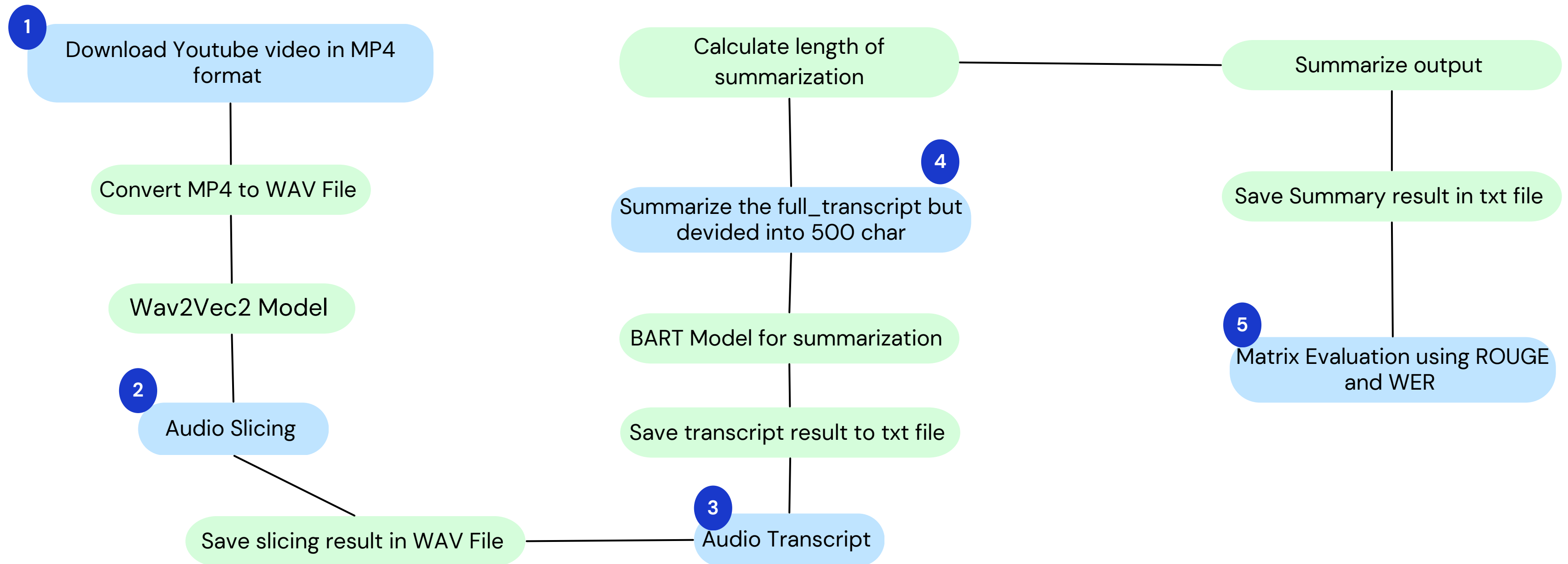


SUMMARIZATION menggunakan BART





Workflow



Hasil

Pada model ini dihasilkan nilai
WER sebesar 26.67%

	Rouge 1	Rouge 2	Rouge L
Recall	47.68%	32.10%	44.96%
Precission	55.66%	33.16%	52.38%
F1-Socre	51.32%	32.62%	48.39%

Thankyou :0

Yee selesai >.<