

Project-III

Part-I

Terms:

1. `Chunk_s`: context length (in seconds)
2. `Stride_s`: stride length (in seconds), higher stride means lesser overlap between consecutive chunks
3. WER: Word error rate, word error between generated transcript and actual transcript
4. Levenshtein and Cosine similarity: character-level and word-level respectively

We explored a very narrow slice of the `{chunk_s, stride_s}` design space for finding a near-optimal transcription.

Chunk_s	Stride_s	WER (%)	Levenshtein (%)	Cosine (%)
30	29	15.71	89.75	96.04
30	28	17.5	88.73	96.27
30	27	22.14	85.56	95.35
30	26	20.71	84.95	96.17
30	25	28.21	81.98	95.43

Below, we highlight repeated phrases in generated transcripts in **red** and spelling mistakes in **blue**. The transcriptions are generated using the script `part1_audio_transcript_openai.py` which use OpenAI's Whisper for transcribing.

Chunk=30, Stride=28

Can you please tell the jury why you're here today? Ms. Hurd accused me of abuse. My ex-husband is suing me. Brutal. Cruel. This is humiliating for any human being to go through. And all false. Amber Hurd forever changed Mr. Depp's life and reputation. Behind the fame, you're going to see who the real Johnny Depp is. Depp was the one who wanted the cameras in midcourt. She didn't. I would argue it's PR. **She did. I would argue it's a PR** campaign disguised as a defamation case. It was the man himself. It's been a social media circus of commentary for VHs and **fluids**. Did you commit any kind of prank? Absolutely not. On my side of the bed was human fecal matter. This has moved away from a news story or a lawsuit. And it's transformed into a cultural moment. **into a cultural moment**. People are live tweeting. People are live streaming. Where does it end? The engagement is phenomenal. Videos can be very easily manipulated and republished. We're being influenced by bots interacting with bots. Johnny Depp is clearly winning right now in the court of public opinion. I've never been so scared in my life. She's acting. This trial is about so much more than Johnny Depp vs Amber. **than Johnny Depp versus Amber** Heard. Why are we all so fascinated with this case? Because they're famous, because of the details. What did the jury see? They just kicked me. It didn't happen. I don't know who to believe. I was kidding. Mr. Depp is your victim, isn't he? This is not so much about the legal merits, but rather what the public perception is. And that leads us to the real question, which is what is the actual **is the actual** truth.

Chunk=30, Stride=29

Can you please tell the jury why you're here today? Ms. Hurd accused me of abuse. My ex-husband is suing me. Brutal. Cruel. This is humiliating for any human being to go through. And all false. Amber Hurd forever changed Mr. Depp's life and reputation. Behind the fame, you're going to see who the real Johnny Depp is. Depp was the one who wanted the cameras in midcourt. She didn't. I would argue it's PR. I would argue it's a PR campaign disguised as a defamation case. It was the man himself. It's been a social media circus of commentary for VHs and fluences. Did you commit any kind of prank? Absolutely not. On my side of the bear was human fecal matter. This has moved away from a news story or a lawsuit. And it's transformed into a cultural moment. People online. People are live tweeting. People are live streaming. Where does it end? The engagement is phenomenal. Videos can be very easily manipulated and republished. We're being influenced by bots interacting with bots. Johnny Depp is clearly winning right now in the court of public opinion. I've never been so scared in my life. She's acting. This trial is about so much more than Johnny Depp versus Amber Heard. Why are we all so fascinated? all so fascinated with this case because they're famous because of the details what did the jury see they just kicked me it didn't happen i don't know who to believe i was kidding sir mr death is your victim isn't he this is not so much about the legal merits but rather what the public perception is and that leads us to the real question which is what is the actual truth Thank you.

We notice increasing stride reduces repetitions but trades off with accurate punctuation and misspelled words. Hence, we adopt {chunk_s=30, stride_s=28} for the next part.

Part-II

Here, we fine-tuned four open-source LLMs from HuggingFace for 10 epochs.

Defamation Analysis on Generated Transcript

Model	Validation Accuracy	Validation F1 score	Defamatory (%)	Most defamatory sentence
google-bert/bert-base-cased	1.0	1.0	78.4	This is humiliating for any human being to go through.
google-bert/bert-large-cased	1.0	1.0	18.9	Videos can be very easily manipulated and republished.
distilbert-base-cased	0.75	0.67	2.7	My ex-husband is suing me.
roberta-large	1.0	1.0	27.0	Where does it end?

Defamation Analysis on Original Transcript

Model	Validation Accuracy	Validation F1 score	Defamatory (%)	Most defamatory sentence
google-bert/bert-base-cased	1.0	1.0	82.9	This is humiliating for any human being to go through.
google-bert/bert-large-cased	1.0	1.0	22.9	Videos can be very easily manipulated and republished.
distilbert-base-cased	0.75	0.67	2.9	My ex-husband is suing me.
roberta-large	1.0	1.0	28.6	And that leaves us to the real question, which is... what is the actual truth?

This demonstrates that the transcript generated in Part-I, which shares very high degree of cosine similarity with the actual transcript results in very similar defamatory detection results as indicated by the fourth column of the above tables.

Packages Required:

librosa, soundfile, transformers [WhisperProcessor, WhisperForConditionalGeneration, AutoTokenizer, AutoModelForSequenceClassification, TrainingArguments, Trainer, EarlyStoppingCallback], torch, numpy, noisereduce, warnings, os, time, pandas, json, sklearn (model_selection/train_test_split), datasets, evaluate, argparse, wandb, nltk

Was wandb API used: No, open source HuggingFace Models were used