**MODEL RESEARCH PART FOR SPEECH RECOGNITION**

MODEL-1 : **Google speech-to-text**

* Accuracy : 85-95%
* Features : Multi-language support, real-time transcription, punctuation handling.
* Strengths : High accuracy, cloud-based, easy integration.
* Weakness : Requires internet, costly for high usage

Git link : <https://github.com/googleapis/java-speech>

MODEL-2 : **Amazon Transcribe**

* Accuracy : 85-93%
* Features : Speaker identification, custom vocabulary, automatic punctuation
* Strengths : Scalable, integrates well with AWS services
* Weakness : Requires AWS setup, occasional lag

Git link : <https://docs.aws.amazon.com/code-library/latest/ug/java_2_transcribe_code_examples.html>

MODEL-3 : **IBM Watson Speech to text**

* Acuuracy : 80-90%
* Features : Customization options, keyword spotting, diarization
* Strengths : Strong customization, supports multiple audio formats
* Weakness : Less accurate than competitors for noisy environments

Git link : <https://github.com/watson-developer-cloud/java-sdk/blob/master/speech-to-text/README.md>

MODEL-4 : **Open AI Wishper**

* Accuracy : 90-96%
* Features : Multilingual, highly robust, work offline
* Strengths : High accuracy, supports various accents and languages
* Weakness : Computationally expensive, slower processing

Git link : <https://github.com/GiviMAD/whisper-jni>

MODEL-5 : **Deepgram**

* Accuracy : 85-95%
* Features : On-premise & cloud deployment, real-time streaming, customization
* Strengths : Fast processing, lower latency, scalable
* Weakness : Less recognized than major competitors

Git link : <https://github.com/deepgram-starters/prerecorded-java-starter>

MODEL-6 : **Vosk**

* Accuracy : 80-90%
* Features : Fully offline, supports multiple platforms, lightweight
* Strengths : No internet required, open-source
* Weakness : Lower accuracy compared to cloud-based solutions

Git link : <https://github.com/topics/vosk-api?l=java>

Conclusion :

Each speech recognition model has its strengths and weaknesses. The best choice depends on the specific use case, whether it's real-time transcription, security needs, or offline capabilities.

* **For enterprise solutions**: Google Speech-to-Text, Amazon Transcribe, and Microsoft Azure Speech are excellent choices.
* **For offline usage**: Whisper and Vosk stand out.
* **For customization**: IBM Watson and Deepgram provide strong flexibility