

Unlocking Value from

Speech Data

Patanjali V, Lead Data Scientist, Tiger Analytics
Anand B, Director Analytics Consulting, Tiger Analytics



EXECUTIVE SUMMARY

Today a lot of unstructured data is being generated in the form of text, images, videos and speech. This data could contain valuable information that companies can utilize to make the right decisions. In this article, we focus on one such form of unstructured data which is speech.

We present a use case, where we analyzed speech in clinical trials to automate a significant part of the operational processes, which has the potential to reduce the quality control costs by half.



What Is Speech Analytics?

Speech has several aspects to it. Some of the elements of speech like words, speech rate, tone, emotions etc. are discernible by humans.

There are other elements that humans don't identify so easily like minor variations in pitch and speech rates.

Speech analytics is the characterization of speech based on these factors to derive actionable business insights from the data.

There are several ways in which speech can be analyzed, based on the type of application:



Full transcription

Full transcription involves conversion of speech into text format in applications like Siri or in transcribing meetings (for example, between a doctor and a patient), conferences, etc. Converting speech into text allows it to be searched more easily.



Speaker diarization

Speaker diarization involves separation of certain sections of speech based on the speaker. While transcribing speech with more than one speaker, like a meeting or a conference, it is important to not just convert speech to text but to identify who the speaker is.



Keyword detection

Keyword detection entails identification of certain specific keywords in an audio. Customer care centers can detect certain keywords like “unhappy” and “disappointed” and use them to monitor agent performance.



Speaker authentication/identification (voice fingerprinting)

Speaker authentication/identification (voice fingerprinting) involves identifying unique characteristics in every speaker’s voice that allow us humans to differentiate between and identify speakers. Some fraud detection applications capture these unique features and create voice fingerprints during customer care interactions and compare against known blacklists.



Emotion detection

Emotion detection involves identification of the emotional state of the speaker. This can help identify irate customers during customer care interactions, among other applications.



Other characteristics of conversation

These are pauses, noise, etc. Characteristics like loud noises or long pauses could be indicators of a bad customer care conversation.

Depending on the type of business problem, the analysis framework would have one or more of the above



Conclusion

Speech analytics is an area with potential applications in almost all businesses that have any form of verbal interaction from call centers to classrooms. With the increase in computing power, and big data technologies, analyzing large volumes of unstructured speech data is becoming increasingly mainstream.

When used appropriately, it can give a company significant reduction in cost as well as strong competitive advantage. Some functions like customer care have started incorporating speech analytics but there is still a long way to go before the full potential is realized.

About the Authors/ Tiger Analytics

Patanjali V, the primary author, is a Lead Data Scientist at Tiger Analytics. He leads advanced analytics engagements that involve complex/unstructured data.

Anand Bharadwaj, the co-author, is a Director at Tiger Analytics. He has 18+ years of experience in the consulting industry and loves to ensure business value realization of analytics solutions.

Tiger Analytics, (www.tigeranalytics.com) provides Big Data and advanced analytics solutions to help businesses make data driven business decisions. We bring deep expertise in data sciences along with understanding of business needs and state-of-the-art technologies to solve business problems.

References

1. http://en.wikipedia.org/wiki/Speech_analytics
2. <https://blog.calltrackingmetrics.com/introducing-keyword-spotting-for-phone-calls/2014/08/>
3. http://en.wikipedia.org/wiki/Speaker_diarisation
4. <http://www1.icsi.berkeley.edu/~vinyals/Files/taslp2011a.pdf>
5. http://en.wikipedia.org/wiki/Mel-frequency_cepstrum
6. <http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs>



Contact Us

info@tigeranalytics.com

www.tigeranalytics.com

+1-408-508-4430

USA - East Coast
2321 Blue Ridge Rd.,
Suite 203
Raleigh, NC 27607

USA - West Coast
4701 Patrick Henry Drive,
Building 16, Suite 14
Santa Clara, CA 95054

India - Chennai
No. D-1, SIDCO Industrial
Estate, Guindy,
Chennai - 600032, India