

Datasheet: Deepgram Model Overview

Deepgram is the most accurate, scalable, reliable and cost effective Automatic Speech Recognition (ASR) on the market. Using our end-to-end deep learning platform organizations can train state-of-the-art AI speech models with their data. Depending on the use case your organization is looking to solve, you will need to attain a certain metric such as level of accuracy or word error rate. Deepgram provides four tiers of models that utilize different training methods to reach your desired goals. The higher your accuracy needs, the more complex model training may be required.

Model Tier	Training Method	Time to Create	Expected Reduction in WER	Description	Use Case
Beginner	N/A	N/A	N/A	Choose from a general, phone call or meeting speech recognition model to transcribe your audio.	<ul style="list-style-type: none"> • Common content across customer base • Used downstream for NLU model creation or call analytics
Intermediate	Auto-Train	Seconds to 1 day	0.5 - 5%	Leverage the latest AI training techniques with intermediate models. From the Deepgram dashboard, self-train a model to recognize audio from your label training data set.	<ul style="list-style-type: none"> • Moderate variations in discussion across customer base • Used downstream for NLU model creation or call analytics
Advanced	AutoML	1 - 7 days	2 - 8%	Powered by AutoML, advanced models perform training tasks such as determining the training curriculum, adjusting width and depth or hyperparameters and performing additional epochs without human intervention.	<ul style="list-style-type: none"> • Moderate to high variations in discussion across customer base • Used downstream for NLU model creation or call analytics
Expert	Deepgram AI Scientist	1 - 3 weeks	+8%	For highly specialized audio use cases where automated training is not enough to reach accuracy requirements, a Deepgram Expert will supervise the model training and tune training to reach the required score.	<ul style="list-style-type: none"> • Many variations in customer conversations, not typical words or sounds • Noisy acoustic environments like call center floors, virtual meetings with multiple speakers, or phone calls with compressed frequencies • New languages • Transcription is customer facing and/or used in real-time such as Conversational AI • Highly regulated industries such as healthcare and financial services

How to Get Started

When using AI enabled ASR, the entire AI speech process to be well architected and aligned to business goals. Once speech AI goals are defined and the use case is clarified, data acquisition, preparation and model training can begin. Deepgram can remove tedious steps from every stage of the speech AI automation process. Our services and software range from data preparation to model training to deployment. Create a [free account](#) to try the Beginner and Intermediate speech models, or contact us to get started with Advanced and Expert models today.