DeepZen

Introdução

1. Captando a mágica da palavra falada

2. Liberando a inovação de IA

3. Chegando antes da concorrência

4. Sobre a DeepZen

Avançando no mercado com as primeiras soluções de inteligência artificial para conteúdo falado convincente com emoções

O mercado de áudio falado está mal atendido devido ao custo proibitivo e ao tempo que leva para encontrar atores talentosos e gravá-los em estúdios. Usando redes neurais desenvolvidas e rolando o IBM® Watson Machine Learning Accelerator, a DeepZen analisa texto e o transforma em fala sintetizada incluindo emoção, lançando sua solução quatro meses antes do planejado.

Desafio de negócio

Até agora, a transmissão efetiva de emoções esteve limitada às vozes humanas. Para atender à demanda por um conteúdo falado convincente, a DeepZen queria oferecer uma alternativa para a gravação em estúdio, que é demorada e de alto custo.

Transformação

A DeepZen desenvolveu o *deep learning* e as redes neurais para reconhecer a emoção no texto e produzir uma fala semelhante à de humanos, com o suporte das soluções IBM® PowerAI que combinam a computação acelerada com o software de inteligência artificial (IA).

Benefícios para os negócios

4 meses
a menos do que o prazo do *time-to-market* esperado, ajudando a DeepZen a ganhar vantagem por ser a pioneira

Oferece
uma fala com emoção aos novos mercados graças à redução de tempo e custo

Rápido
crescimento da DeepZen devido ao suporte da IBM para atingir clientes potenciais

Introdução

1. Captando a mágica da palavra falada

2. Liberando a inovação de IA

3. Chegando antes da concorrência

4. Sobre a DeepZen

Captando a mágica da palavra falada

Somente os seres humanos possuem a capacidade de comunicar emoções por meio da fala. Desde as propagandas de televisão até os audiolivros, os atores inspiram empatia e prendem a atenção das pessoas ao expressar as palavras com sentimento.

No entanto, a gravação de vozes é um processo trabalhoso e caro. Existe uma escassez de estúdios e atores talentosos. A produção de um audiolivro típico custa milhares de dólares e demora semanas. Isso gerou uma enorme demanda não atendida por conteúdo falado entre o grande público com deficiência visual, dislexia ou que simplesmente gostam de ouvir a fala gravada. Além disso, mais empresas estão usando o conteúdo falado para cativar os clientes nos canais digitais, aumentando ainda mais a demanda.

A DeepZen foi estabelecida para preencher essa lacuna no mercado. A empresa foi estruturada para combinar técnicas de conversão de texto para voz e de processamento de linguagem natural para fornecer uma alternativa mais eficiente em relação a custos e tempo do que a gravação em estúdio.

Taylan Kamis, CEO e cofundador da DeepZen, explica: “Nosso objetivo não é eliminar o trabalho dos dubladores, mas resolver os problemas de capacidade no mercado atual. Identificamos as emoções no texto automaticamente e usamos amostras de voz, cujos royalties pagamos aos dubladores, combinadas com a tecnologia de síntese de fala para produzir áudios de voz convincentes. Para fazer isso, é necessário criar redes neurais grandes e complexas. Isso exige uma quantidade enorme de energia de processamento para produzir resultados precisos rapidamente, portanto, precisávamos da plataforma de tecnologia certa para concretizar nossa visão”.

> “Mesmo antes de começarmos a compartilhar qualquer coisa com a equipe IBM, dava para notar que eles estavam tão empolgados quanto nós com o nosso conceito e, desde então, o nível de comprometimento conosco nunca caiu.”
>
> — Taylan Kamis, CEO e cofundador da DeepZen

Introdução

1. Captando a mágica da palavra falada

2. Liberando a inovação de IA

3. Chegando antes da concorrência

4. Sobre a DeepZen

Liberando a inovação de IA

A DeepZen escolheu a plataforma IBM PowerAI Enterprise para dar suporte às suas soluções de produção de IA para audiolivros e locuções.

As soluções combinam os servidores IBM Power® Systems AC922 que apresentam os processadores IBM POWER9™ mais recentes, estruturas populares de *deep learning* e ferramentas de IA para desenvolvimento eficiente, oferecendo à empresa um ambiente abrangente para inovação de IA.

Kerem Sozugecer, CTO da DeepZen, relembra: “Descobrimos a tecnologia em uma reunião sobre o IBM PowerAI que aconteceu em Londres. No evento, conversamos com engenheiros da IBM e da NVIDIA, que explicaram os recursos exclusivos da plataforma e como ela é projetada para acelerar a criação de soluções de IA.”

A DeepZen participou do programa “Startup with IBM”, em seguida, adotou o Programa IBM Global Entrepreneur, fornecendo acesso aos créditos de tecnologia que usou para experimentar os recursos na IBM Cloud.

Logo, a empresa estava pronta para comprar a plataforma IBM PowerAI Enterprise e incentivou o parceiro comercial Platinum da IBM Meridian IT a hospedar a solução em seu data center.

Ao apresentar o Large Model Support, a solução IBM permite um desempenho muito alto, até mesmo para as tarefas de *machine learning* mais amplas e complexas. A DeepZen pode alternar entre várias estruturas de *deep learning*, todas otimizadas para serem usadas no IBM POWER Architecture®.

“O IBM PowerAI consegue lidar com modelos muito grandes, o que nos oferece liberdade para experimentar”, comenta Sozugecer. “Também podemos alternar entre as estruturas TensorFlow, Caffe e PyTorch ou usá-las em paralelo. Esse nível de flexibilidade é um ativador real de inovação”.

Introdução

1. Captando a mágica da palavra falada

2. Liberando a inovação de IA

3. Chegando antes da concorrência

4. Sobre a DeepZen

Chegando antes da concorrência

Com as ferramentas de desenvolvimento de IA e uma enorme quantidade de energia de processamento em mãos, a DeepZen ultrapassou sua programação de desenvolvimento original. Como resultado, a empresa tem a vantagem por ser a pioneira em um mercado que apresenta grande potencial.

“O IBM PowerAI Enterprise acelerou significativamente nosso tempo de treinamento de modelo”, diz Kamis. “Quatro meses antes da meta e estamos prontos para entrar em produção, podendo começar a aproveitar a enorme oportunidade que nos espera.”

Ao permitir a produção de audiolivros e locuções por uma fração de custo e tempo de execução dos métodos tradicionais, a DeepZen colocará um conteúdo falado convincente ao alcance dos novos clientes. Os usuários podem gerar vozes diretamente do texto e editá-las usando as ferramentas da empresa, contando com uma flexibilidade sem precedentes.

“Estamos revolucionando o setor de audiolivros e locução, sem substituir a função dos dubladores”, comenta Kamis. “Com o suporte da tecnologia IBM, podemos capacitar os usuários a refinar o áudio de voz até que ele fique perfeito ou criar várias versões com diferentes sotaques ou idiomas, por exemplo”.

Trabalhando com a IBM, a DeepZen está preparada para a expansão global. A empresa pode dimensionar facilmente suas soluções IBM PowerAI e está aproveitando os recursos da IBM para lançar sua oferta revolucionária no mercado.

Kamis concluiu: “Desenvolvemos nossas soluções para a perfeita integração com o processamento de linguagem natural do IBM Watson, o que abre a porta para os clientes Watson existentes. Mesmo antes de termos começado a compartilhar o código com a equipe IBM, dava para notar que eles estavam tão empolgados quanto nós com o nosso conceito e, desde então, o nível de comprometimento conosco nunca caiu. Juntos, estamos criando uma fala semelhante à humana, que é convincente, barata e de rápida produção”.

> “O IBM PowerAI Enterprise acelerou significativamente nosso tempo de treinamento de modelo. Quatro meses antes da meta e estamos prontos para entrar em produção, podendo começar a aproveitar a enorme oportunidade que nos espera.”
>
> — Taylan Kamis, CEO e cofundador da DeepZen

Introdução

1. Captando a mágica da palavra falada

2. Liberando a inovação de IA

3. Chegando antes da concorrência

4. Sobre a DeepZen

Sobre a DeepZen

A Deep Zen Ltd. (DeepZen) está criando soluções específicas para os setores que produzem falas com qualidade semelhante à humana e repleta de emoções. A empresa tem o objetivo de ajudar os clientes a evitar o custo e o tempo das sessões de gravação em estúdios, além de fornecer a eles a flexibilidade para editar a fala usando ferramentas intuitivas. A DeepZen está sediada em Londres, Reino Unido.

Solução composta por:

- IBM® PowerAI Enterprise
- IBM Power® Systems AC922
- IBM POWER9™

Dê o próximo passo

Para saber como as soluções de infraestrutura IBM Power podem ajudar o seu negócio, entre em contato com o especialista aqui.