



Phirio

IA - analyse, génération de sons : WaveNet, AWS Transcribe

IA050

Durée: 2 jours

1 610 €

29 au 30 janvier
29 au 30 mai

9 au 10 septembre
20 au 21 novembre

Public :

Développeurs, data-scientists, chefs de projet, et toute personne souhaitant connaître les possibilités de l'IA pour l'analyse, la détection, ou la génération de sons.

Objectifs :

Comprendre les principes de fonctionnement, et savoir quels sont les outils existants pour le traitement ou la génération de sons.

Connaissances préalables nécessaires :

Connaissances de base du machine learning.

Programme :

Introduction

Définition d'un son, bases de traitement de signal numérique
Différents types de son : voix, musique, bruits (industriels, d'ambiance, ..)
Technologies mise en oeuvre : traitement du signal, machine learning, deep learning (réseaux de neurones)
Intervention de l'IA à différentes étapes du processus de traitement du son :
enregistrement, filtrage, montage, tri d'enregistrements, analyse, génération, etc ...

Applications

Traitement de la voix, reconnaissance vocale, compréhension et analyse de discours ou de commandes vocales, traductions, retranscriptions, domaine médical, etc ...
Analyse de musiques (classement, proposition, reconnaissance style, compositeurs, etc ..)
Détection de bruits, filtrage, sécurité, suppression de bruits (amélioration productions musicales)
Analyse et séparation de sources sonores
Génération de musique, synthèse vocale



Phirio

Outils

MetaVoice : modification de la voix
WellSaid, Play.ht : transformation de texte en audio
Séparation de pistes audio : AudioStrip
Synthèse vocale : wavenet
Audio et vidéo : Sonix, NVIDIA Jarvis, IBM Watson,
Traitement d'enregistrements audio : Amazon Transcribe (retranscription en texte, analyse de contenu, statistiques de mots clés, reconnaissance d'interlocuteur, ...)

Atelier : démonstration de 'Text to Speech' d'IBM Watson

Retranscription de discours en temps réel avec Amazon Transcribe

IA génératives

Principe de fonctionnement
GAN (Generative Adversarial Network)
Exemples: MusicLM, Boomy, Mubert, Soundraw, Beatoven, ..
Google MusicLM : génération de musique à partir d'une description en texte
Mubert : génération de musique libre de droits,
description en texte, choix du style, ..

Atelier : création de morceaux de musique avec Mubert

Evolutions

Limites : problèmes éthiques, détournement de voix, clonage de voix,
Sources : données utilisées en apprentissage
Aspects juridiques, droits d'auteurs, liens avec l'industrie musicale