

Etude de marché

Outil d'augmentation de la donnée

Abou Assaf Mahyr-Florian
Ilan' Daumont-Ouk
Ethan Chemla
Alexandre Ngo

INTRODUCTION

SOMMAIRE

1 - Définir le marché

2 - Evaluer la demande

3 - Identifier les offres existantes

4 - Analyse des informations

1 - Définir le marché

Objectifs

- Réaliser une photographie générale du marché
- S'assurer de la cohérence générale du projet
- Constituer un outil support pour le financeur
- Valider l'offre commerciale avec le marché cible

Identification et évolutions

Plusieurs questions :

- Évolution du marché?
- Secteurs du marché ciblés?
- Qui sont les clients/utilisateurs?

Produits concurrents

- Rechercher les produits qui seront des concurrents directs/indirects
- Découvrir les nouvelles tendances du marché
- Identifier des opportunités

Les acteurs

Identifier et définir les principaux acteurs de votre marché:

- les futurs concurrents
- les entreprises
- les clients
- les utilisateurs

2 - Evaluer la demande

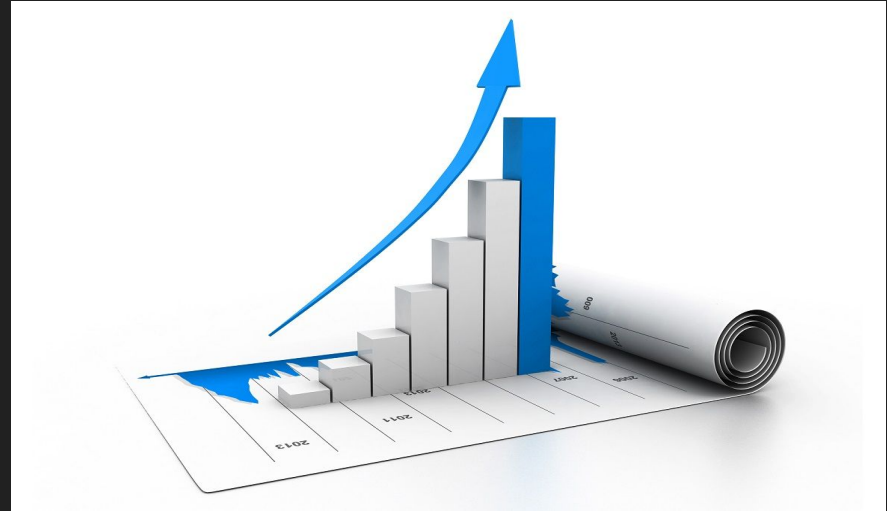
Les différentes cibles utilisant l'augmentation de la donnée

- Le secteur de l'augmentation de la donnée étant en expansion, de nombreuses zones géographiques sont concernées : l'Amérique du Nord, l'Europe, l'Asie et le Moyen-Orient.
- Différents secteurs sont sollicités par l'augmentation de la donnée, les deux secteurs principaux sont le secteur médical et automobile

Les besoins et les motivations

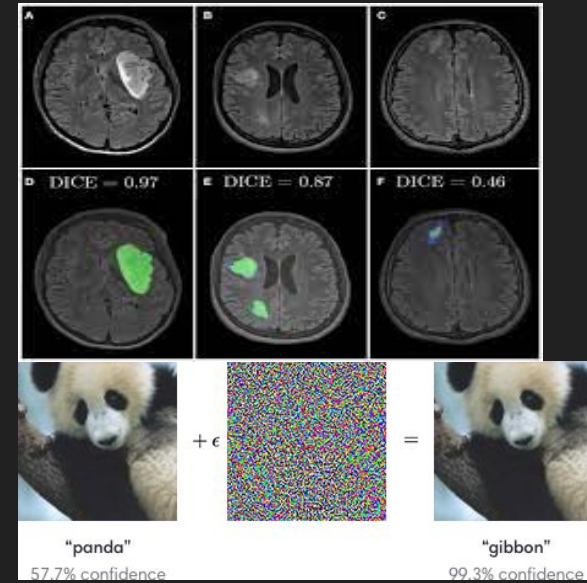
- L'une des opportunités d'utilisation de l'augmentation de la donnée est **d'éviter une collecte de données supplémentaires** lorsqu'il y a un manque de données via **l'amélioration des modèles d'IA**.

Cela mène à partir de données réelles à l'obtention de résultats valorisés avec une méthode **peu coûteuse et efficace**.



- Imagerie : - GAN : Generative Adversarial Networks

- Apprentissage automatique contradictoire



- Musique : Tempo / Pitch
- D'autres secteurs : le secteur spatial, militaire, les pompiers ...

sources : <https://neovision.fr/data-augmentation-solutions-manque-donnees/>
<https://www.v7labs.com/blog/data-augmentation-guide#h3>

3 - Identifier les offres existantes



- *IBM* a beaucoup investi dans le domaine du Big Data.
- L'entreprise a déployé des logiciels permettant de gérer de grands jeux de données à l'aide d'applications comme Cognos ou SPSS .
- Elle fournit également une plateforme cloud pour le stockage de données avec de nombreuses fonctionnalités tel que le deep learning et l'analyse des données.
- Enfin, *IBM* propose des outils d'analyse prédictive, des modèles pré-construits, des services de conseil et des systèmes d'intelligence artificielle. Ces outils peuvent être utilisés pour enrichir les données existantes et pour fournir des informations précises et complètes sur le marché.



- Boite de consulting en IA, fournit une solution adaptée aux clients dans tous les domaines de l'IA et également dans l'**augmentation de données**.
- Utilise le principe de **transformations de données**, et celui de **simulation et données de synthèse**.
- Nettoyage et standardisation des jeux de données.
- Utilisations de **moteurs 3D** et de **réseau génératifs** afin de créer des environnements virtuels réalistes.
- Emploi des **Generative Adversarial Networks**, technique de génération de données réalistes basée sur une compétition entre deux réseaux de neurones.

Google BigQuery :



Google
BigQuery

- Spécialiser dans l'analyse des données et la datascience
-
- **Google Knowledge API**, requêtes SQL pour augmentation du jeu de donnée dans un tableau BigQuery
- **Entity Reconciliation API**, regroupement des données, extraction des données pour construire des graphiques représentatif de l'ensemble des données.



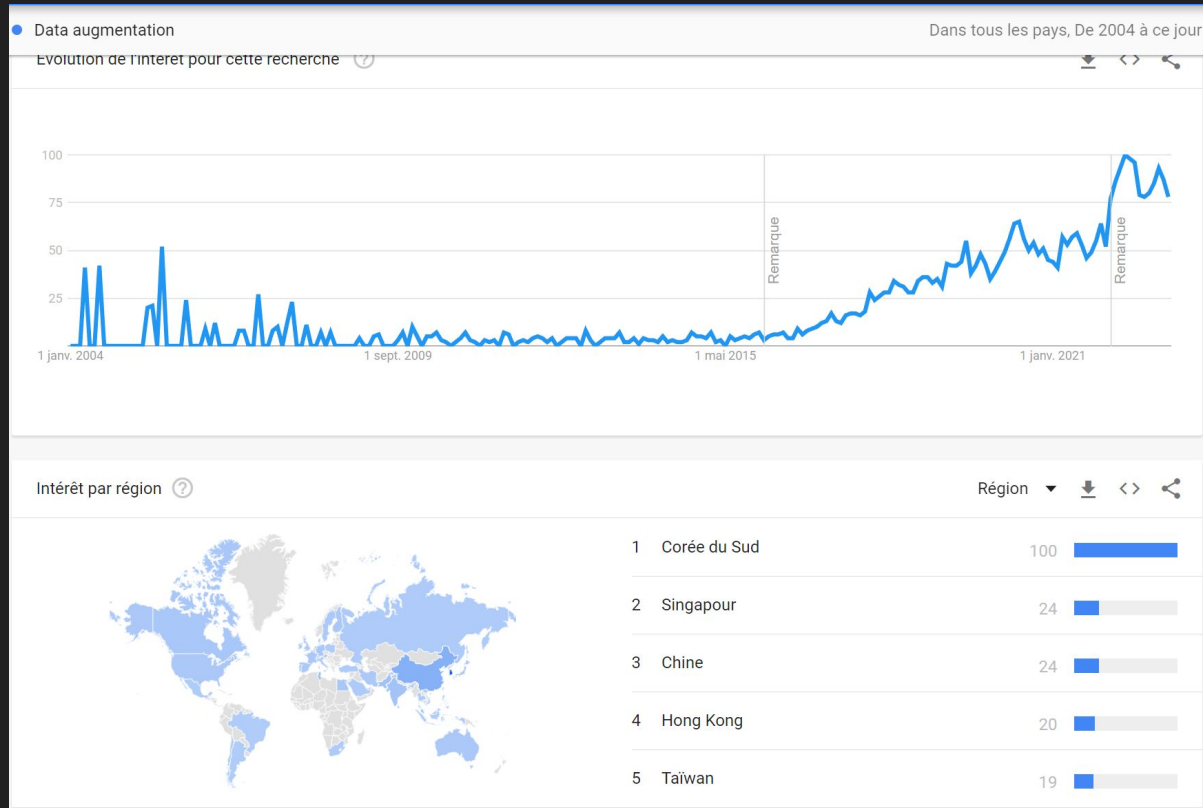
Pour résumer :

- ❑ Parmi les principales entreprises proposant des services d'augmentation des données pour leurs clients figurent *IBM, HP, Google, SAP et Oracle*.
- ❑ Ces offres incluent des outils d'analyse prédictive, des modèles pré-construits, des services de conseil et des systèmes d'intelligence artificielle.
- ❑ Elles utilisent de **nombreux algorithmes** afin de procéder à l'augmentation de données et ces entreprises permettent une **analyse complète** des données fournies.
- ❑ Divers fournisseurs de services cloud, dont *Oracle*, proposent des solutions d'**analyse et d'intégration de données**, ainsi que des solutions de visualisation et de reporting pour aider les organisations à tirer le meilleur parti de leurs données.
- ❑ De plus, des sociétés telles qu' *IBM* et *Microsoft* proposent des solutions spécialisées pour la **collecte et l'organisation des données**.

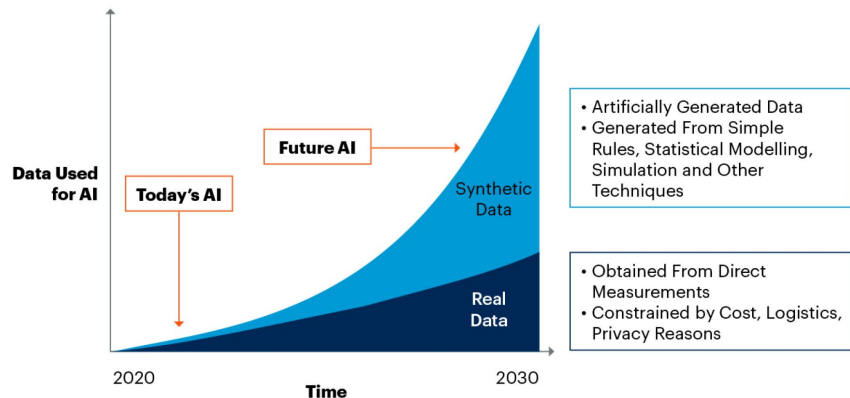
4 - Analyse

- L'augmentation de la donnée : technologie récente qui se développe depuis la **création du machine learning en 1959**. Plusieurs modifications d'images possibles(rembourrage,rotation,bruit,contraste,translation)
- Démocratisation par son aspect stratégique : **Manque accru de données** pour alimenter les projets d'IA. Cela permet au modèle de gagner en **qualité**.

Nombre de recherches du terme “Data Augmentation”



By 2030, Synthetic Data Will Completely Overshadow Real Data in AI Models

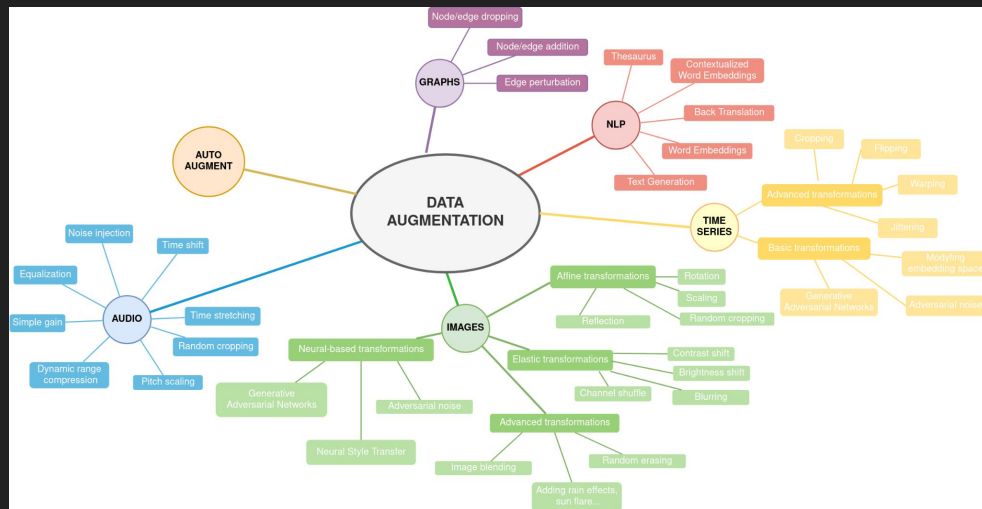


Source: Gartner
750175_C

Gartner

La **donnée synthétique** sera de plus en plus utilisée dans les années à venir d'après Gartner ...

L'augmentation de la donnée touche plusieurs domaines ...



Le marché

- Peu d'entreprises proposant ces services, mais plutôt des **outils** d'augmentation de la donnée (AWS, IBM Watson, Azure Machine Learning, Google Cloud...). Également des librairies open source (Tensorflow, PyTorch...)
- Cibles : secteur **industriel**, la **finance**, le médical, et toute autre firme ayant des services informatiques utilisant du machine learning.

Tableau des revenus et des environnements utilisés pour l'augmentation de la donnée par certaines grandes entreprises

Entreprise	Revenus	Environnement utilisé
Microsoft	\$ 21 Billion	<i>Azure Machine Learning</i>
IBM	\$ 17 Billion	<i>IBM Watson</i>
Intel	\$ 10 Million	<i>OpenCV</i>
Open source	\$ 5 Million	<i>PyTorch</i>

Conclusion