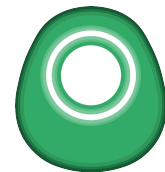


Picsell.ia

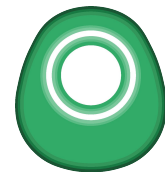
# Rapport d'avancée

Picsell.ia - 17/06/2019



Picsell.ia

# Contextualisation



Picsell.ai

Rendez-vous pris en mars 2019, nous vous avons présenté notre projet de start-up :

Nous sommes spécialisée en annotation semi automatique de base **de données d'images, de vidéos**, de textes et de sons **nécessaires au déploiement de vos solutions IA.**

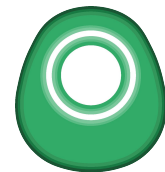
Nous avons besoin de données pour tester notre service, vous nous avez donc envoyé votre dataset contenant des déchets **PET non PET / Alimentaire non Alimentaire / Bouteille Barquette.**

Ce dataset s'inscrivant dans **votre volonté de développer des solutions de tri via intelligence artificielle**, la labellisation de donnée n'étant pas votre principal problème car vous avez la main sur vos données, vous nous aviez proposé de vous présenter nos résultats afin que vous puissiez déterminer si notre service répond ou non à vos besoins.

Nous avons toutefois tenu à faire de **nombreuses analyse** de vos données, afin de vous donner le plus d'informations possible et vous présenter les **pistes possibles de développement de votre projet.**

Dans un second temps nous allons vous présenter notre service et ce qu'il peut vous apporter.

# Les prérequis de votre projet



Picsell.ia

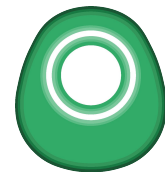
## **Pertinence de vos données pour votre use-case:**

- Dataset ayant la même proportion d'image dans chaque classe.
- Qualité suffisante des images pour obtenir un process de qualité

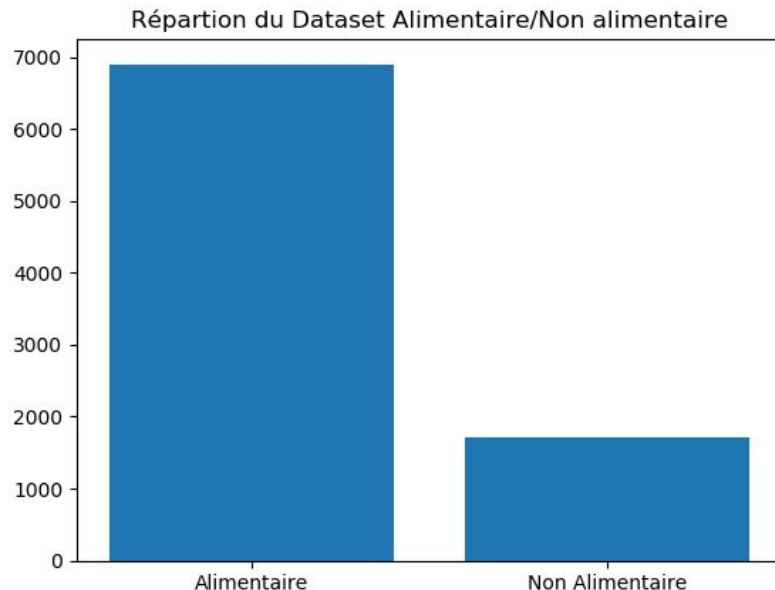
## **Faisabilité de votre use-case:**

- Bonnes performances de prédiction si l'on a un dataset labellisé

# Prérequis 1 : Proportion du dataset

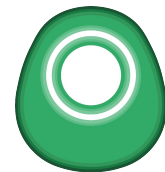


Picsell.ia

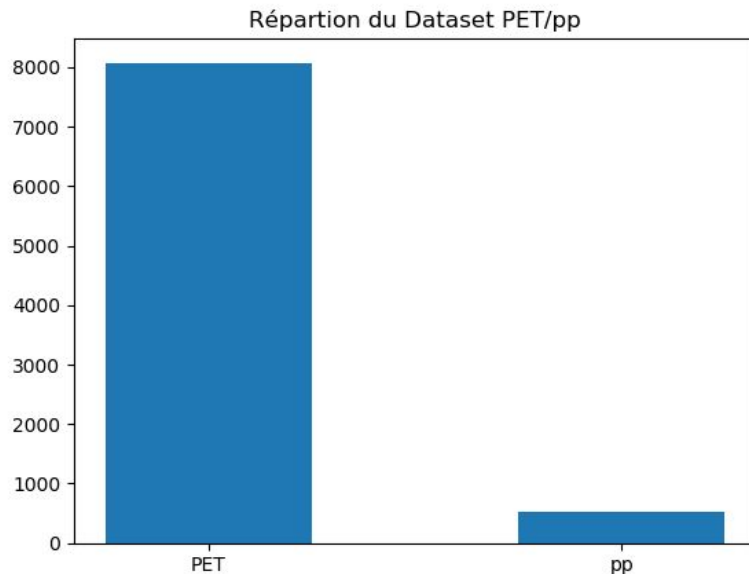


On voit ici que nos deux **classes**, représentant respectivement les déchets **alimentaires et non alimentaires** ne sont **pas représentées en proportion égale** dans le dataset. Ainsi lors de l'apprentissage de vos algorithmes, vous allez instaurer un **biais de prédiction**. C'est à dire que votre IA va interpréter qu'il y a plus de déchets alimentaires présents sur vos tapis roulant, ainsi il va préférer reconnaître le déchet comme alimentaire quand ce dernier à un "doute".

# Prérequis 1 : Proportion du dataset

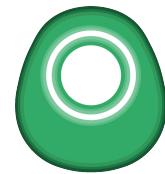


Picsell.ia

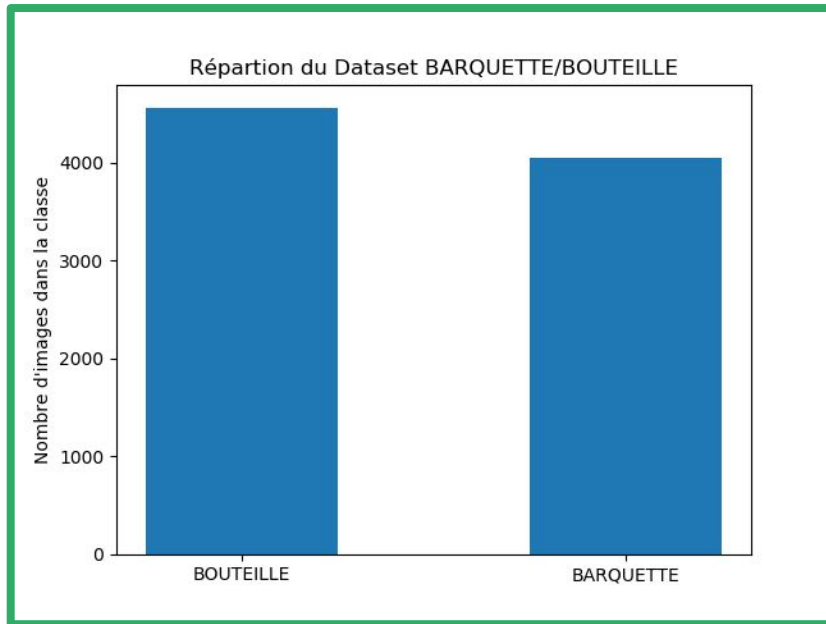


Nous pouvons faire la **même remarque** par rapport aux **classes PET/pp**.  
Notons aussi que le **phénomène est exacerbé**. En l'état votre dataset ne permettra pas de prédire la composition de vos déchets. L'algorithme se contentera de prédire que l'image est un déchet PET.

# Prérequis 1 : Proportion du dataset

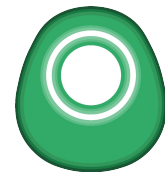


Picsell.ia



Ici, on note une **qualité du dataset**, en effet les **classes bouteille et barquette** sont représentées dans la **même proportion**. Le **prérequis** pour ce use-case ci est donc **rempli**.

# Prérequis 2 : Qualité des images



Picsell.ia



*Alimentaire*

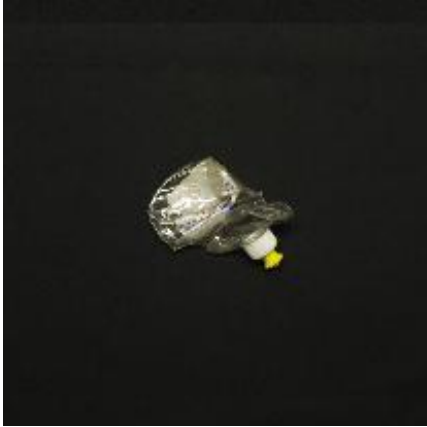


*Non - Alimentaire*

Nous entendons par **qualité** des images **deux aspects** :

- Le premier étant la qualité en terme de **résolutions**, ici vos images sont en **207x207 pixels**, pour votre use-case l'étude des **textures étant très pertinentes**, la résolution semble trop faible.
- Le deuxième étant la qualité au sens de la **perception humaine**. Un bon moyen de le représenter serait de se poser la question suivante : “ Suis-je capable de différencier ces deux images à l'oeil nu ?” Ici la réponse semble être non, cela veut dire que vous avez besoin d'informations non contenu dans des images classiques. Vous aviez mentionné des **images hyperspectrale**, peut-être serait-ce plus recommandé.

# Prérequis 2 : Qualité des images



*PET*



*pp*

Ici nous pouvons faire les **même remarques** que précédemment. Néanmoins nous aimerions **insister** sur la **résolution des images**, en effet différencier des matériaux passent **impérativement par la détection de textures**. Ainsi comme dit antérieurement, la résolution de vos images est un point clef pour reconnaître ces types de déchets.



# Prérequis 2 : Qualité des images



*PET*



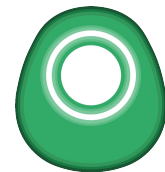
*pp*

Ici, la **qualité des images semble suffisante** pour votre use-case. En effet en terme de résolution, cette dernière est suffisante pour distinguer les **couleurs** présentent sur les bouteilles, ainsi il est **possible de reconnaître ces deux classes**.

En terme de qualité perceptuelle aussi, les bouteilles, bien que écrasées semble toujours avoir **une forme reconnaissable**.

**Ainsi ce use case remplit le prérequis.**

# Prérequis 3 : Faisabilité



Picsell.ia

Les deux use-cases PET/PP et A/NA étant écartés par les prérequis précédent non remplis nous allons maintenant étudier la faisabilité de votre troisième use-case:

→ **BOUTEILLE/BARQUETTE**

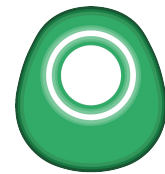
Afin de tester la faisabilité, nous avons fait tourner un réseau de neurones convolutif à 5 couches.

Sans optimiser les algorithmes nous arrivons tout de même à **prédire la bonne classe 85.7 % du temps.**

Ainsi ces résultats qui ne sont pas les meilleurs résultats possibles car nous le répétons, nous n'avons pas optimisé les algorithmes, sont **très encourageants**.

**Le prérequis est donc rempli.**

# Notre service



Picsell.ia

Nous fournissons un service d'**annotation d'image et de conseil en intelligence artificielle**.

En couplant intelligemment l'**action humaine de labellisation et nos algorithmes**, nous optimisons la préparation de votre base de donnée et vous donnons une insight pertinente sur les algorithmes à utiliser pour votre use-case.

Pour vous donner un ordre d'idée, nous labellisons : **1 000 images en 25 minutes**.

Ainsi pour créer une database susceptible d'obtenir des bonnes performances pour discriminer Bouteille/Barquette il vous faut un dataset d'**image différentes** ( Sans "data augmentation" → cf **rotation de vos objets lors de la prise de photo** ) d'environ **5 000 images par classe**. Notre service peut vous fournir une telle base en 7h de travail comprenant :

- 4 h de labellisation assisté par nos algorithmes
- 1 h de process de data augmentation pour vous fournir une base prête à l'emploi
- 2 h d'analyse et de rédaction de rapport sur vos données et conseils adaptés à vos compétences en interne

Nous vous permettrons donc d'**accélérer la phase de recherche et développement de vos projets d'intelligence artificielle** en vous donnant toutes les clefs pour la **réussite de votre projet**.