Mankouri Jalil |   |



Table des matières

[Perspective d’évolution 1](#_Toc508623573)

[Architecture du système 2](#_Toc508623574)

[I- Google Cloud Platform 2](#_Toc508623575)

[1) Cloud Datalab 3](#_Toc508623576)

[2) Cloud Dataproc 4](#_Toc508623577)

[3) BigQuery 5](#_Toc508623578)

[II- TensorFlow et ses réseaux de neurones 7](#_Toc508623579)

[1) L’inférence 7](#_Toc508623580)

[2) Fonction de coût 7](#_Toc508623581)

[3) Entrainement 7](#_Toc508623582)

[4) TensorBoard : Visualisation 7](#_Toc508623583)

[Solutions 7](#_Toc508623584)

[I- Nos données pour l’apprentissage 7](#_Toc508623585)

[II- Optimisation du réseau de neurones 7](#_Toc508623586)

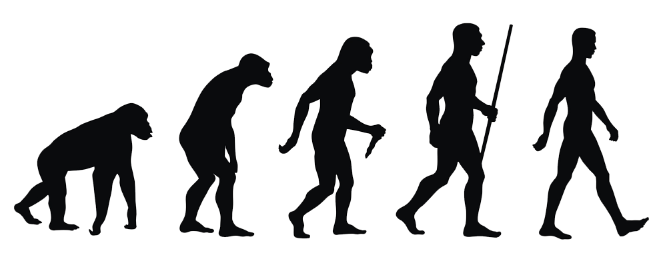
[1) Nombre de neurone et de couche 7](#_Toc508623587)

[2) Paramètres de TensorFlow 7](#_Toc508623588)

[III- Résultats 7](#_Toc508623589)

[Bibliographie 7](#_Toc508623590)

# Perspective d’évolution



L’objectif premier de cette perspective d’évolution de notre projet serait d’analyser toute sorte d’image de nos utilisateurs automatiquement.

Lister le contenu présent sur chaque image pourrait nous permettre d’affiner le profil de nos utilisateurs au fur et à mesure qu’ils utilisent notre application.



Tout cela afin de leurs proposer ensuite un meilleur contenu à découvrir, donc de toucher la dimension marketing de STYLBUM en fidélisant les utilisateurs par le biais d’une excellente expérience utilisateur.

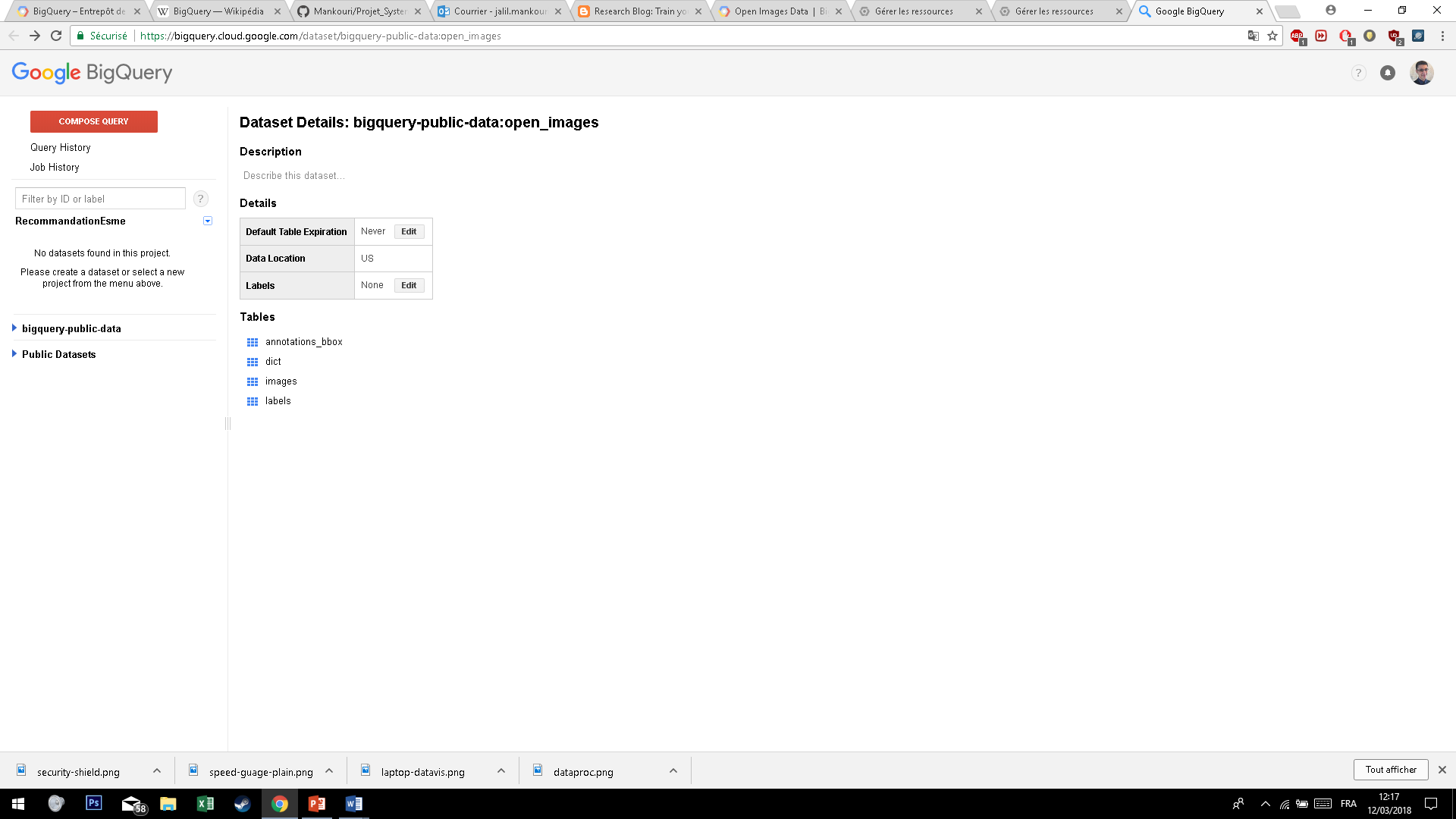


# Architecture du système

## Google Cloud Platform



Google Cloud Platform est une des solutions fournies par Google de Cloud Computing. Elle propose un hébergement sur la même infrastructure que celle utilisée par son moteur de recherche. Google Cloud Platform est donc une plateforme fournissant aux développeurs des services permettant de construire plus aisément des sites web comme des applications complexes comme le Deep Learning.



Basé sur le cloud, cette plateforme met à disposition une multitude de services modulaires comme le stockage d’informations, le calcul, des applications de traduction de texte et prévision, etc…

Notre choix c’est porté sur Google Cloud Platform plutôt qu’AWS d’Amazon pour ses prix concurrentiel, sa facilité d’utilisation, la performance qu’elle fournit malgré la gratuité de sa solution et ses services que nous détaillerons plus bas.

### Cloud Datalab



Cloud Datalab est une solution Open Source interactive permettant d’analyser, parcourir, transformer et visualiser des données en plus de fournir la possibilité d’élaborer des modèles de Machine Learning depuis la plateforme principale.



Cloud Datalab a pour but de facilité la compréhension de nos analyses en matière de science des données. Il se combine parfaitement avec BigQuery, Cloud Storage et des programmes Python.

Ce qui nous a particulièrement intéressé était sa capacité à créer des modèles de Machine Learning à partir de nos données pour faire des prédictions tout en les évaluant et optimisant à l’aide de TensorFlow.

### Cloud Dataproc



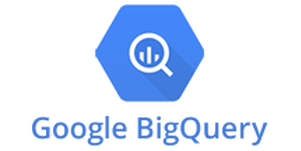
Afin de traiter rapidement de grand set de données, nous avons de plus souvent besoin de recourir aux services d’Apache comme Hadoop, Spark, Pig et Hive.

Cloud Dataproc nous apporte ses solutions afin de créer rapidement des clusters de tailles variable, désactivable quand le besoin ne se fait plus ressentir afin de réduire les coûts.

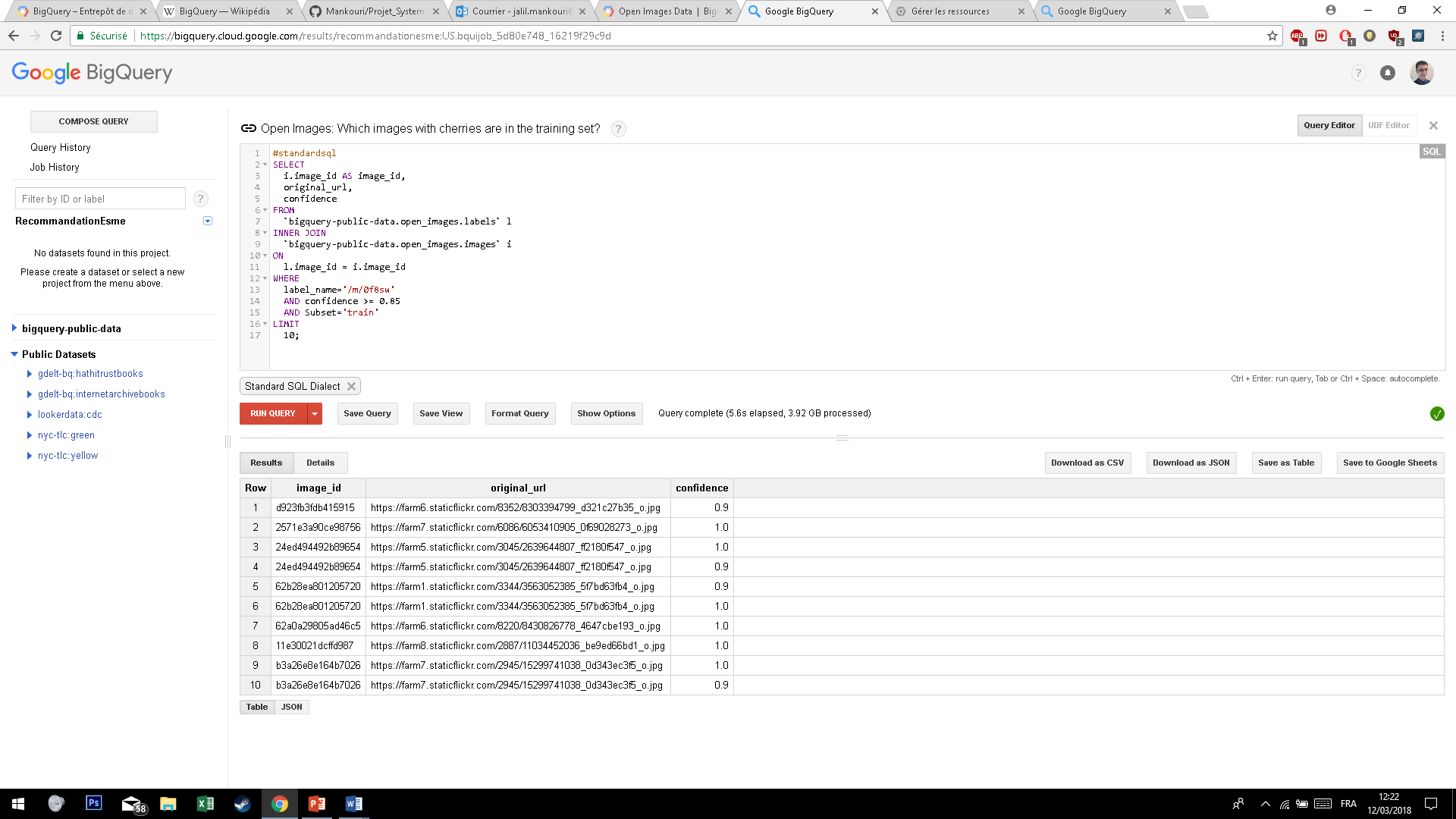


Ce qui nous a tout de suite plu dans les services que proposait Google Cloud Platform était que chaque outil que nous implémentions dans notre solution était complémentaire d’un autre. Par exemple, il est pour nous envisageable que nous utilisions, en association de BigQuery, Hive sur Hadoop. Ou bien si nous voulions créer des pipelines de transformation de données à l’aide de Spark, nous pourrions utiliser un autre service appelé Cloud Dataflow.

### BigQuery



BigQuery nous permet d’analyser jusqu’à des pétaoctets de données, tout cela sans avoir aucune infrastructure à gérer étant donné qu’il fonctionne sans serveur. Cela nous permet donc de nous concentrer sur le traitement de nos données afin d’en dégager les informations dont nous avons besoin par le biais de requêtes SQL puissante et optimiser pour de grand ensemble.



BigQuery est très utilisé dans le domaine du Big Data, qu’il s’agisse de start-up ou d’entreprises présentent dans le CAC40.



Ce que nous aimons par-dessus tout est l’externalisation de la sécurisation de nos services, en effet Google nous protège des menaces extérieures en sécurisant nos applications en chiffrant et répliquant nos données sur plusieurs datacenters pour nous offrir une sécurité, disponibilité et durabilité maximale.

## TensorFlow et ses réseaux de neurones

blabla

### L’inférence

### Fonction de coût

### Entrainement

dz

### TensorBoard : Visualisation

# Solutions

## Nos données pour l’apprentissage

## Optimisation du réseau de neurones

### Nombre de neurone et de couche

### Paramètres de TensorFlow

## Résultats

# Bibliographie

**Aucune source spécifiée dans le document actif.**