A close up of a logo

Description generated with very high confidence

|  |
| --- |
| CapGemini Invent |
| Rapport de projet : Catégorisez automatiquement des questions |
| Open Classroom |

|  |
| --- |
| François Lemeille  27/09/2018 |

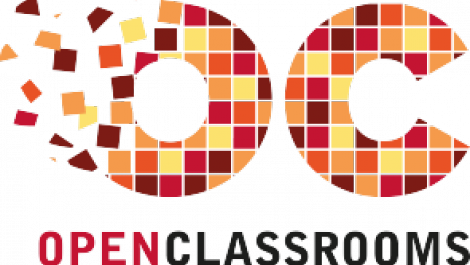


Table des matières

[Table des matières 1](#_Toc525827120)

[I. Introduction 2](#_Toc525827121)

[1. Contexte du projet 2](#_Toc525827122)

[2. Aperçu des données 2](#_Toc525827123)

[II. Nettoyage et premières analyses 3](#_Toc525827124)

[1. Corps et titre de la question 3](#_Toc525827125)

[2. Tags 3](#_Toc525827126)

[3. Premières analyses 3](#_Toc525827127)

[III. Méthode « Semi-supervisée » 9](#_Toc525827128)

[1. Approche 9](#_Toc525827129)

[2. Construction d’une métrique 9](#_Toc525827130)

[3. Résultats 9](#_Toc525827131)

[IV. Méthode supervisée 9](#_Toc525827132)

[1. Approche 9](#_Toc525827133)

[2. Résultats 9](#_Toc525827134)

[3. Approfondissements 9](#_Toc525827135)

[V. Conclusion 9](#_Toc525827136)

# Introduction

## Contexte du projet

Dans le cadre de ce cinquième projet, il nous est demandé de créer un outil de proposition automatique de tags à partir de questions. Les questions considérées sont des questions provenant de StackOverflow, un célèbre forum d’échange dans le domaine de la programmation. Les questions portent donc sur des problèmes de développement informatique et c’est la communauté qui apportent les réponses. Un système de vote a également été mis en place afin d’identifier les questions et les réponses jugées les plus utiles, les plus claires et les mieux formulées.

StackOverflow, réelle référence du domaine, propose également d’associer des tags lorsque à la question posée par l’utilisateur pour une meilleure identification et classification des différentes questions. L’utilisation des tags nécessite une certaine connaissance du domaine de la programmation et de l’utilisation de StackOverflow. C’est dans ce but que nous avons développé notre outil de suggestion de tags : pour aider les utilisateurs novices à tagger leur question à partir des propositions de notre outil.

## Aperçu des données

# Nettoyage et premières analyses

HdqHGLKJFQD

## Corps et titre de la question

A mon arrivée, après 2 journées d’accueil et de formation j’ai très rapidement été envoyé en mission. J’ai rejoins z

### Equipe centrale

* , le Comité d’Orientation et le Comité de pilotage opérationnel. Ces comités ont lieu trime

.

## Tags

Malgré l’intérêt que je portais à cette mission et le fait de monter en compétences sur de nombreux sujets, j’ai demandé auprès de mon management de pouvoir changer de mission. En effet, comme écrit plus haut, je souhaitais durant ce stage pouvoir intervenir sur des sujets plus innovants. C’était également une des discussions développées lors de mes entretiens et j’ai donc réexprimé ces désirs aux personnes en charge de mon stage courant Novembre.

J’ai été entendu, puisque une semaine plus tard, j’ai été remplacé sur mon poste par un autre

## Premières analyses

### Contexte et objectifs de la mission

Le

à

Le processus que je devais automatiser était un processus de contrôle de déversement de deux primes d’assurances « P11 » et « P13 » comme nommées au sein de Cardif. Il fallait comparer les données indiquées en comptabilité avec les données indiquées en contrôle de gestion pour s’assurer que celles-ci correspondaient et que l’on observait aucun écart entre les deux. Pour cela il faut aller collecter les données dans différentes applications internes à Cardif. Cette partie de collecte a finalement été sortie du périmètre du POC car elle nécessite un long temps de Bilan du stage et projet professionnel

# Méthode « Semi-supervisée »

## Approche

## Construction d’une métrique

## Résultats

# Méthode supervisée

## Approche

## Résultats

## Approfondissements

# Conclusion