

# Etude de l'opinion et de la confiance sur TripAdvisor.

Faiza Belbachir\*

\*Institut polytechnique des sciences avancées (IPSA)  
phdups@gmail.com

## 1 Introduction

De nos jours, avec l'explosion du Web la majorité des services, produits, films actuels sont marqués, jugés et appréciés par un grand nombre d'utilisateurs. Ces informations s'avèrent très utiles pour la prise de décision. Il existe différentes plateformes qui se sont spécialisées sur la notation d'avis (système d'Etoiles) comme Amazon, Airbnb, Booking, Tripadvisor. Leur principe est d'aider l'utilisateur à faire son choix en fonction de l'avis et de l'expérience d'autres utilisateurs. L'enjeu de ces sites est donc de pouvoir déterminer au mieux les commentaires les plus fiables et qui répondent aux attentes des internautes. Il est difficile de déterminer ou de mesurer cette fiabilité ou plus exactement il est difficile d'attribuer une confiance à un commentaire. Dans ce travail, nous nous sommes intéressés à répondre à cette problématique.

## 2 Approche proposée

Afin de déterminer les documents porteurs d'opinion fiables, nous procédons selon les étapes suivantes :

- Déterminer les documents porteurs d'opinion ;
- Déterminer les caractéristiques qui permettent de mesurer la confiance ;
- Calculer le score de confiance ;
- Combiner entre l'opinion et la confiance pour déterminer les documents porteurs d'opinion fiable.

Pour calculer le score d'opinion  $ScoreOp(d)$  (equation 1) nous utilisons une approche lexicale basée sur le lexique SentiWordnet. Pour mesurer la confiance du commentaire nous nous sommes basés sur des caractéristiques dont certaines sont spécifiques à la plat-forme d'avis considérée TripAdvisor : le nombre de commentaires (NC) ; le nombre de jugements utiles (NUC) : lorsque vous lisez un commentaire, vous pouvez le juger utile ou pas et donc, plus un utilisateur a de jugements utiles plus il est considéré comme digne de confiance ; moyenne de commentaires utiles (AN) : il peut être intéressant de connaître la moyenne des commentaires jugés utiles. En effet, si un utilisateur commente beaucoup mais a peu de commentaires utiles il est moins digne de confiance que celui qui en commente moins mais avec beaucoup de commentaires jugés utiles par d'autres. Nous allons donc prendre en considération le ration entre le nombre de commentaires utiles et le nombre de commentaires comme caractéristique de confiance ; nombre de badges (NB) et le nombre de points (TP) : ces scores sont attribués

Etude de l'opinion et de la confiance sur TripAdvisor.

par TripAdvisor en fonction de différentes caractéristiques. Plus le nombre de badges est élevé plus l'utilisateur est expérimenté; durée de service (D) : le nombre d'années d'abonnement sur le site est une caractéristique importante. Pour mesurer le score de confiance nous sommes les valeurs des différentes caractéristiques (equation 2).

$$ScoreOp(d) = \sum_{si \in sens(w)} \frac{(Neg(si) + Pos(si))}{|sens(w)|} \quad (1)$$

$$ScoreTC(d) = NC + NUC + AN + TP + NB + D \quad (2)$$

Pour retourner des documents de confiance et porteurs d'opinion nous combinons linéairement le score de confiance avec le score d'opinion dans l'equation suivante :

$$ScoreTS(d) = \lambda * ScoreOp(d) + (1 - \lambda) * ScoreTC(d) \quad (3)$$

Avec  $\lambda$  est un coefficient qui varie selon la valeur donnée.

Nos expérimentations ont été faites sur la collection Tripadvisor qui contient 200. 000 hôtels et de 1800000 utilisateurs extraite des travaux des auteurs Wang et al. (2010).

Nous analysons l'impact du score de confiance TC sur le score final TS (voir le graphe 1)

Nous observons que plus le Score TC est fort plus le Score TS est élevé et inversement plus le

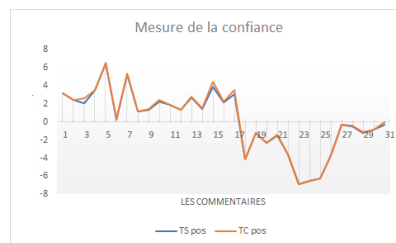


FIG. 1 – Le score TC et TS.

Score TC est faible et plus le Score TS est faible. Nous concluons ainsi que plus on a confiance à l'utilisateur plus son score final est élevé.

### 3 Conclusion

Nous introduisons la notion de confiance du commentaire. L'expérimentation a été faite sur la collection TripAdvisor. L'analyse des résultats prouve que les informations qui entourent le commentaire jouent un rôle sur la fiabilité/confiance du commentaires sur la collection TripAdvisor.

### Références

Wang, H., Y. Lu, et C. Zhai (2010). Latent aspect rating analysis on review text data : A rating regression approach. In *Proceedings of the 16th ACM SIGKDD, KDD '10*, NY, USA, pp. 783–792. ACM.