

# Propositions pour améliorer une méthode de prédiction du succès d'une campagne de financement participatif

Alexandre Blansché et Xavier Mazur

LORIA (UMR 7503), Campus Scientifique B.P. 239  
54506 Vandœuvre-lès-Nancy Cedex, France  
alexandre.blansche@univ-lorraine.fr

**Résumé.** Le financement participatif est un mode de financement d'un projet faisant appel à un grand nombre de personnes qui a connu une forte croissance avec l'émergence d'Internet et des réseaux sociaux. Cependant plus de 60 % des projets ne sont pas financés, il est donc important de bien préparer sa campagne de financement. De plus, en cours de campagne, il est crucial d'avoir une estimation rapide de son succès afin de pouvoir réagir rapidement (restructuration, communication) : des outils de prédiction sont alors indispensables. Nous proposons dans cet article plusieurs pistes d'amélioration pour la prédiction du montant levé lors d'une campagne de financement participatif en utilisant l'algorithme  $k$ -NN. La première proposition consiste à utiliser un algorithme de *clustering* afin de segmenter l'ensemble d'apprentissage et faciliter le passage à l'échelle. La seconde proposition consiste à extraire des caractéristiques pertinentes depuis les séries temporelles et les informations sur les campagnes pour avoir une représentation vectorielle.

## 1 Introduction

En quelques années seulement, le financement participatif a connu une forte croissance. Ce phénomène émergent est encore peu étudié, mais soulève de nombreuses questions dans divers champs scientifiques. L'une des principales interrogations est de connaître à l'avance le succès d'une campagne de financement participatif par des techniques de prévisions.

Nous proposons dans cet article de reprendre la méthode de prédiction du montant final levé lors d'une campagne de financement participatif utilisant l'algorithme  $k$ -NN proposée dans Blansché et al. (2017), et d'y apporter deux améliorations afin de faciliter le passage à l'échelle sur des données de plus grande taille et d'accroître les performances des prédictions. La première amélioration proposée consiste à découper l'ensemble d'apprentissage en groupes homogènes à l'aide d'un algorithme de classification non supervisée. La recherche des  $k$  plus proches voisins pourra alors se limiter à un faible nombre de *clusters* plutôt qu'à l'ensemble d'apprentissage complet.