Classificação de projeto do Kickstarter

Diogo Barbosa, Lucas Zacarias, Luiz Felipe, Luiz Reis, Victor Martins

# 1 Introdução

## 1.1 Kickstarter

Kickstarter é o maior site de financiamento coletivo do mundo e que busca apoiar projetos inovadores. Mais de 1 bilhão de dólares já foram doados a projetos através da sua plataforma.

## 1.2 Objetivo

Temos como objetivo neste estudo entender a influência das características de um projeto para aprovação de seu financiamento coletivo.

## 1.3 KPI

Temos como KPI o número de investidores do projeto e o valor arrecadado do projeto.

# 2 Análise Exploratória da base de dados

## 2.1 Base de dados

A base de dados escolhida foi retirada do kaggle os dados originais possuem 5 classes, sendo elas: sucesso (57,7%), falha (34,4%), cancelados (4,3%), em andamento (3,4%) e suspensos (0,2%); para fins do experimento estamos interessados apenas nas classes sucesso e falha; analisando a base de dados notamos também a presença de outliers (que correspondiam a % da base), a saber, projetos que tinham como meta de arrecadação centavos e constavam na classe sucesso com valor arrecadado muito maior do que pediram (x%) e outros que não tinham investidores, mas tinham dinheiro investido (y%).

## 2.2 Relação dos atributos com a classe alvo

## 2.3 Correlação dos atributos

# 3 Pré-processamento da base de dados

## 3.1 Criação de novas features

Em relação às features vetorizamos o campo “nome” eliminando stopwords

## 3.2 Normalização

# 4 Modelagem

## 4.1 Curva ROC

## 4.2 Matriz de confusão

## 

## 4.3 Algoritmo 1

## 4.4 Algoritmo 2

## 4.5 Algoritmo 3

## 5. Seleção do melhor modelo

## 5.1 Algoritmos e parâmetros

## 

## 6. Conclusão

## 7. Referências