

# Relatório Técnico: Implementação e Análise do Algoritmo de Regressão Linear

Mateus Tamarindo Gonzaga

Rafael Ottoni Rodrigues Gonçalves

## **RESUMO**

O objetivo deste projeto é desenvolver um modelo preditivo usando o algoritmo de Regressão Linear para resolver um problema de inferência sobre taxa de engajamento dos principais influenciadores do instagram.

### **Sobre o conjunto de dados**

O Instagram é um serviço de rede social americano de compartilhamento de fotos e vídeos fundado em 2010 por Kevin Systrom e Mike Krieger, e posteriormente adquirido pelo Facebook Inc.. O aplicativo permite que os usuários carreguem mídia que pode ser editada com filtros e organizada por hashtags e marcação geográfica. As postagens podem ser compartilhadas publicamente ou com seguidores pré-aprovados. Os usuários podem navegar pelo conteúdo de outros usuários por marcação e localização, visualizar o conteúdo em alta, curtir fotos e seguir outros usuários para adicionar seu conteúdo a um feed pessoal.

A rede Instagram é muito usada para influenciar pessoas (os seguidores dos usuários) de uma maneira específica para uma questão específica - o que pode impactar a ordem de algumas maneiras.

### **SOBRE O ARQUIVO**

O conjunto de dados pode ser baixado pelo link [Top Instagram Influencers data](#), disponibilizado pelo Kaggle. Neste arquivo,

basicamente há 10 atributos. Ele foi ordenado com base na classificação que foi decidida com base em "seguidores".

**rank:** Classificação do Influenciador com base no número de seguidores que ele tem

**channel\_info:** Nome de usuário do Instagrammer

**influence score:** Pontuação de influência dos usuários. É calculado com base em menções, importância e popularidade

**posts:** Número de posts que eles fizeram até agora

**followers:** Número de seguidores do usuário

**avg\_likes:** Média de curtidas em posts do Instagrammer (total de curtidas/total de posts)

**60\_day\_eng\_rate:** Taxa de engajamento dos últimos 60 dias do Instagrammer como fração de engajamentos que eles fizeram até agora

**new\_post\_avg\_like:** Média de curtidas que eles têm em novos posts

**total Likes:** Total de curtidas que o usuário obteve em seus posts. (em bilhões)

**country:** País ou região de origem do usuário.