

Análise de sentimentos de avaliações de aplicativos da Google Play Store usando AWS Comprehend

Cristiano O. Ribeiro; Rafael Givanildo da Silva

¹Departamento de computação - Universidade Federal de São Carlos
São Carlos - SP - Brasil

***Resumo.** Este documento descreve os conceitos e metodologias utilizados no segundo trabalho da disciplina de Inteligência Artificial. Neste projeto, foram utilizadas técnicas de Aprendizado de Máquina (AM) para a extração de sentimentos dos textos de avaliação de aplicativos em inglês disponíveis na Google Play Store.*

1. Introdução

Atualmente, nós desenvolvedores, estamos cada vez mais conectados com os usuários finais de nossas aplicações, o que torna a troca de informações mais rápida e eficiente. No entanto, muitas vezes os dados disponíveis são extremamente vastos, fazendo com que seja quase impossível que atendamos adequadamente a toda a demanda. Um grande exemplo deste cenário está nas avaliações recebidas por nossos aplicativos nas lojas virtuais, este rico recurso nem sempre é totalmente utilizado, pois a quantidade de dados é tão vasta que muitas vezes a extração de informação se torna bastante problemática. Nesse sentido, este projeto aplica análise de sentimentos nos textos de avaliação de aplicativos em inglês disponíveis na Google Play Store, facilitando assim a análise da percepção do usuário para com as funcionalidades oferecidas pelo aplicativo.

2. Revisão Conceitual

Antes de introduzir o que foi implementado neste artigo, é interessante se pontuar sobre alguns conceitos importantes utilizados neste trabalho.

2.1. Linguagem Natural

Linguagem Natural é qualquer linguagem desenvolvida naturalmente pelos seres humanos, de forma não intencional. Essas linguagens tem uma sintaxe e uma gramática. Um desafio para lidar com essa Linguagem Natural se dá devido a alta ambiguidade existente nas suas sentenças.

2.2. Linguagens Formais

Linguagens Formais são linguagens que não tem nenhuma ambiguidade e são muito utilizadas em computação.

2.3. Processamento de Linguagem Natural

Processamento de Linguagem Natural ou PLN é um conjunto de teorias/técnicas para resolução de problemas computacionais que dependam do processamento da linguagem humana - fala ou escrita. As pesquisas em PLN abrangem três grandes aspectos da língua natural: o som (fonologia), a estrutura (morfologia e sintaxe) e o significado (semântica e pragmática).

3. Tecnologias Utilizadas

Para a realização deste projeto, foram utilizadas as seguintes tecnologias:

3.1. AWS Comprehend

O AWS Comprehend é um serviço de processamento de linguagem natural (PLN) que utiliza aprendizado de máquina para encontrar conhecimentos e relacionamentos em textos. Este serviço oferece diversas funcionalidades, dentre elas, extração de palavras-chave, identificação de idioma e extração de sentimentos. Este projeto foca apenas na extração de sentimentos, uma técnica que consiste em extrair informações de textos em linguagem natural, identificando a polaridade dos mesmos.

3.2. Google Play Store

Google Play Store é a loja oficial de aplicativos da Google para o sistema operacional Android. A loja permite que usuários baixem e desenvolvam aplicativos Android, além de também oferecer músicas, filmes, revistas, livros e outros conteúdos multimídia. Este projeto foca apenas nos textos de avaliações escritos por usuários da Google Play Store.

3.3. Python

Python é uma linguagem de auto-nível amplamente utilizada no mercado de trabalho e na academia. Esta linguagem foi escolhida para este projeto devido à sua enorme aplicabilidade em problemas de PLN, além da excelente documentação e integração com os serviços da AWS.

4. Metodologia

A metodologia empregada nesse projeto consiste na extração de avaliações da Google Play Store sobre um determinado aplicativo informado pelo usuário. Tais avaliações são devidamente parseadas e enviadas ao AWS Comprehend para extração de pontuações classificando a polaridade de cada avaliação. No caso de várias avaliações realizadas em um mesmo dia, é calculada a média entre tais pontuações. Finalmente, as informações são apresentadas ao usuário na forma de um gráfico (Figura 1) em que o eixo X contém as datas das avaliações e o eixo Y contém a intensidade do sentimento contido na avaliação. Assim, o desenvolvedor é capaz de acompanhar diariamente a percepção dos usuários para com as funcionalidades oferecidas pelo aplicativo.

5. Resultados

Os resultados apresentados pelo programa se mostraram bastante fiéis à percepção de um humano. O AWS Comprehend foi capaz de interpretar os sentimentos dos textos de maneira correta e eficiente, como podemos ver nos exemplos a seguir (Figura 2, Figura 3 e Figura 4) criados utilizando a interface web do serviço.

6. Conclusões

A partir da análise realizada neste projeto, vimos o quanto é simples a implementação de programas que utilizam técnicas de Aprendizado de Máquina e o quão vantajoso essas técnicas podem ser em determinados cenários. Como próximos trabalhos, podemos tentar extrair e utilizar mais informações disponíveis no AWS Comprehend, além de implementarmos nossos próprios algoritmos de AM, vindo a melhorar cada vez mais os resultados obtidos.

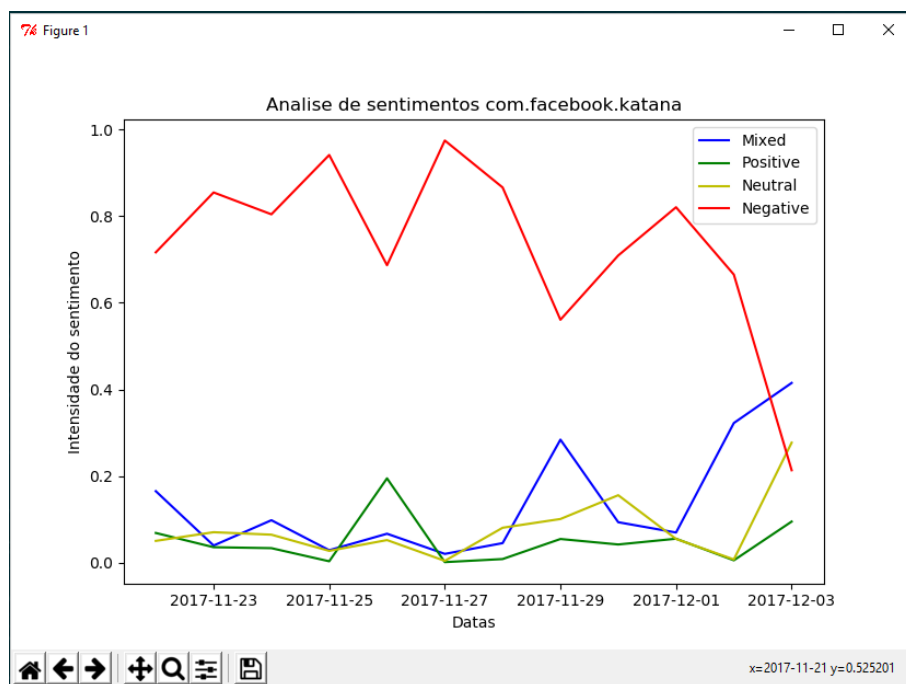


Figura 1. Avaliação do App do Facebook

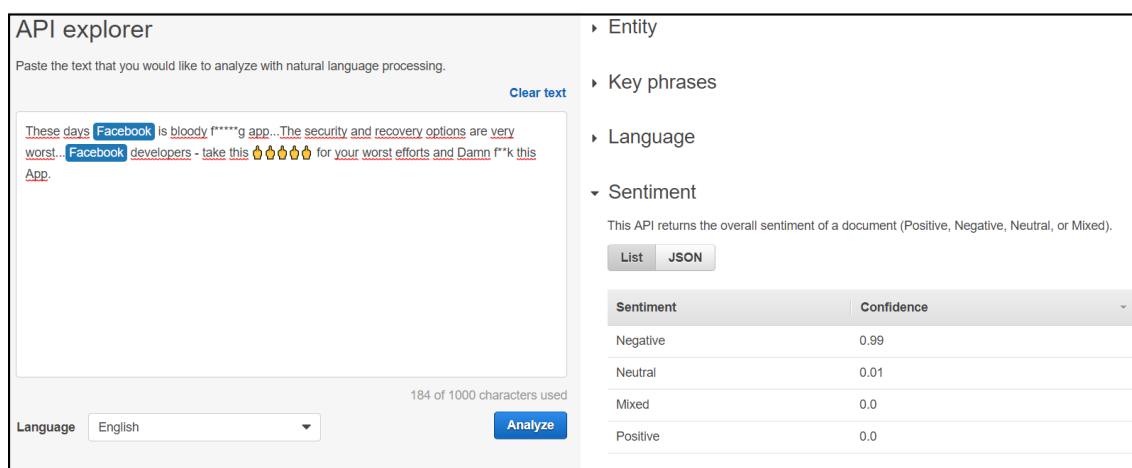


Figura 2. Análise de texto negativo

7. Referencias

AWS Comprehend - <https://aws.amazon.com/comprehend/>
Acesso em 03 de Dezembro de 2017

Google Play Store - <https://play.google.com/store/apps?hl=en>
Acesso em 03 de Dezembro de 2017

Stack Overflow - <https://stackoverflow.com/a/38978341>
Acesso em 03 de Dezembro de 2017

IA-pln - <https://www.ime.usp.br/~slago/IA-pln.pdf>
Acesso em 12 de Dezembro de 2017

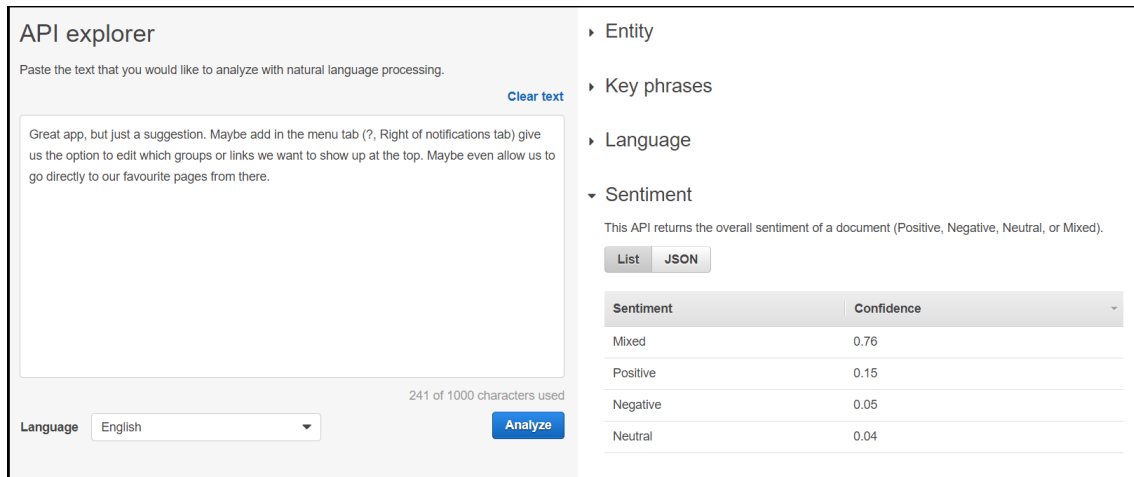


Figura 3. Análise de texto misto

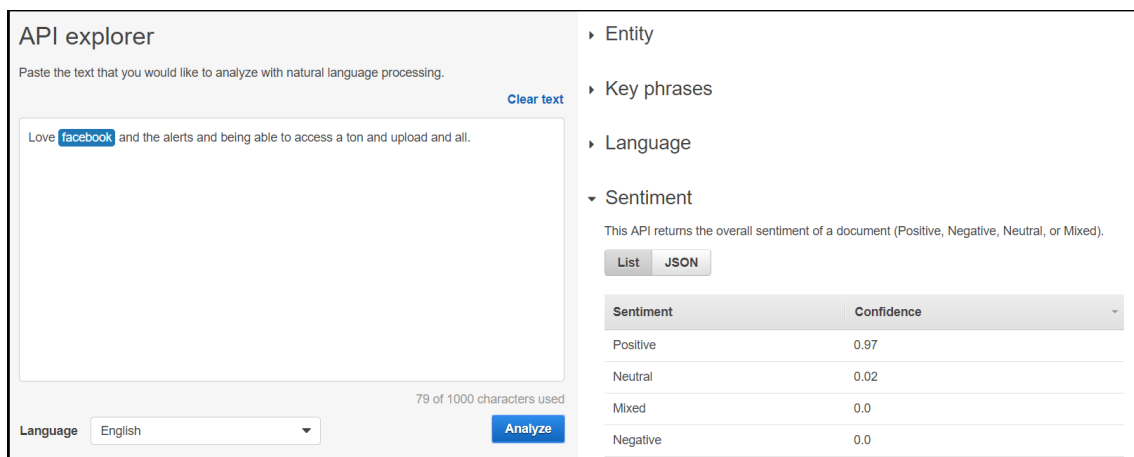


Figura 4. Analise de texto positivo