

## SENTEDU: APLICAÇÃO MÓVEL PARA ANÁLISE DE SENTIMENTO NA EDUCAÇÃO

Daniela Santos Silva<sup>(1)</sup>; Gilson Pereira dos Santos Júnior<sup>(2)</sup>; Thiers Garretti Ramos Sousa<sup>(3)</sup>;  
Lauro Barreto Fontes<sup>(4)</sup>; George Leite Júnior<sup>(5)</sup>.

<sup>(1)</sup>Graduada em Sistemas de Informação, Discente egresso do Instituto Federal de Sergipe, ;<sup>(2)</sup>Mestre em Ciência da Computação, Docente do Instituto Federal de Sergipe, gilson.pereira@ifs.edu.br; <sup>(3)</sup>Mestre em Ciência da Computação, Docente do Instituto Federal de Sergipe, thier\_sl82@hotmail.com; <sup>(4)</sup>Mestre em Ciência da Computação, Docente do Instituto Federal de Sergipe, lauro.fontes@ifs.edu.br; <sup>(5)</sup> Mestre em Ciência da Computação, Docente do Instituto Federal de Sergipe, george.junior@ifs.edu.br.

**Resumo:** Conhecer a percepção dos discentes, docentes e da comunidade geral sobre diferentes aspectos da Instituição tais como: infraestrutura, gestão e qualidade de ensino é fundamental para futuras tomadas de decisão dos gestores. É possível mensurar a imagem que a comunidade tem sobre a Instituição através da análise de sentimentos das postagens em redes sociais. Neste contexto, o objetivo deste trabalho é apresentar o SentEdu, uma aplicação móvel para analisar o sentimento da comunidade sobre a educação no IFS, por meio das postagens na página do Facebook da Instituição. Os resultados gerados mostraram-se positivos, visto que a ferramenta obteve uma precisão de 60% de acerto na análise dos comentários.

**Palavras-Chave:** Redes Sociais, SenticNet, IFS, Análise de Sentimentos

**Abstract:** Knowing the students, teachers and the general community's perception about different aspects of the Institution, such as: infrastructure, management and the quality of the teaching is fundamental for future decision making of the managers. It's possible to measure the impression the community has about the Institution through the analysis of feeling of posts in social media. In this context, the purpose of this research is to present the SentEdu, a mobile application used to analyse the community's feeling about the education in IFS, through posts on the Institution's

Facebook page. The results obtained are positive, once the tool had a precision of 60% of correctness in the analysis of the comments.

**Keywords:** Social Networks, SenticNet, IFS, Sentiment Analysis

### INTRODUÇÃO

As instituições de ensino buscam, a cada dia, melhorar o serviço prestado, otimizar o uso de recursos, engajar e motivar os alunos para maximizar o aprendizado, visando a excelência no ensino. Entretanto, a imagem de uma instituição de ensino não se limita às quatro paredes de uma sala de aula, pois envolve vários aspectos como: o conforto e a acessibilidade da escola, a infraestrutura, a forma como é administrada e a disponibilidade de recursos para atividades, dentre outros.

Com advento das mídias sociais provenientes da Web 2.0, surgiu uma nova possibilidade de interação entre as pessoas que forneceu mecanismos para a livre publicação de conteúdos e a exposição de opiniões, contribuindo para construção de grandes repositórios, principalmente em estruturas no formato de textos (PANG; LEE, 2008). O volume de dados publicados na Web vem crescendo exponencialmente e produzindo enorme quantidade de dados que devem ser filtrados e analisados, pois podem conter informações ocultas que são relevantes para diferentes domínios como: educação, indústria, ciências sociais, economia, dentre outros.

Este cenário proporcionou o surgimento de uma nova área de pesquisa denominada mineração de opinião, ou análise de sentimentos, cujo objetivo é desenvolver sistemas para analisar as opiniões, as avaliações, as atitudes e as emoções das pessoas em relação à certa entidade (LIU, 2007). A literatura aponta alguns desafios nesta área, sendo que grande parte das dificuldades dizem respeito à linguagem dos textos avaliados, à detecção de ironia, à dificuldade de tratar a subjetividade dos textos e a impossibilidade de analisar automaticamente dados em português (MOREIRA *et al.*, 2016).

Na educação, a análise de sentimentos se apresenta como um mecanismo não intrusivo para conhecer a percepção da comunidade (alunos, pais, técnicos e professores) sobre diferentes aspectos da instituição: infraestrutura, gestão, qualidade de ensino, demanda por novos cursos, eventos culturais, pesquisa e extensão, políticas contra retenção e evasão, dentre outros. Segundo Moreira *et al.* (2016) a análise das postagens dos discentes servem como insumo para que professores e tutores trabalhem de forma a melhorar a condução das aulas e à disseminação do conteúdo disponibilizado, visando atrair e manter a atenção dos alunos.

Diante do contexto, este trabalho apresenta o SentEdu, uma aplicação móvel para analisar o sentimento da comunidade sobre a educação no IFS, por meio das postagens na página do Facebook da Instituição. Assim, O SentEdu permitirá aos gestores do IFS analisar, periodicamente, o sentimento da comunidade para embasar tomada de decisões em setores como ensino, pesquisa, extensão, infraestrutura ou administração.

## TRABALHOS RELACIONADOS

A análise automática de sentimentos vem sendo utilizada para extrair das postagens em redes sociais a percepção dos usuários sob diferentes aspectos nos últimos anos como, por exemplo, nos trabalhos de (MALHEIROS *et al.* 2013),

(TEDESCHI AND BENEDETTO 2015), (MOREIRA *et al.* 2016) sumarizados na Tabela 1.

**Tabela 1** - Comparação com trabalhos relacionados

Trabalho	Fonte de Dados	Processamento	Sentimento
(MOREIRA <i>et al.</i> 2016)	Facebook	N/A	SenticNet, SentiStrength, LIWC
(MALHEIROS <i>et al.</i> 2013)	Facebook	N/A	SenticNet
(TEDESCHI AND BENEDETTO 2015)	Twitter	N/A	SentiWordNet
Trabalho Proposto	Facebook	N/A	SenticNet

Malheiros *et al.* (2013) apresentaram uma ferramenta de análise de sentimentos para mensagens compartilhadas no Facebook utilizando o SenticNet como base de conhecimento. Os experimentos realizados com a ferramenta mostraram que a abordagem consegue classificar mensagens adequadamente e suporta a grande quantidade e velocidade na qual elas são enviadas para as redes sociais.

Já em 2015, Tedeschi e Benedetto (2015) desenvolveram uma aplicação de monitoramento de marcas e análise de postagens de mídias sociais na nuvem, baseada na análise de sentimentos, denominado Social Brand Monitoring (SBM). Um crawling coleta postagens no servidor do Twitter e um algoritmo de análise de sentimento utilizando o SentiWordNet, na língua inglesa, os classifica em 7 (sete) níveis de polaridade (Strong Positive, Positive, Weak Positive, Neutral, Weak Negative, Negative, Strong Negative). Além do nível de polaridade do comentário, os autores mensuram a popularidade de quem postou, a fim de compreender a influência do autor na opinião dos seus seguidores.

Recentemente, Moreira *et al.* (2016) utili-

zaram os dados de interações de alunos no Facebook para identificar os sentimentos envolvidos em cada comentário. Após coleta dos textos, o sentimento dos mesmos foi analisado de forma manual e automática com as técnicas SenticNet, SentiStrength e LIWC. Os resultados obtidos foram comparados com a avaliação realizada pelos autores dos comentários, a fim de identificar qual método mais se aproximava dos reais. As ferramentas utilizadas na comparação realizada indicaram diferenças entre os resultados obtidos e o real: enquanto as ferramentas utilizadas classificaram mais de 40% dos comentários como neutros, a análise dos autores das mensagens indicou que 71% dos comentários eram positivos. Na classificação foram identificados outliers e divergências na categorização da intensidade dos sentimentos por cada método. Nenhuma abordagem, incluindo a realizada pelo pesquisador, foi considerada suficientemente eficiente, uma vez que o maior nível de acurácia obtido foi inferior a 70%.

## SENTEDU

O SentEdu é uma aplicação para dispositivos móveis com sistema operacional Android, construído na linguagem Java, capaz de coletar mensagens de páginas da rede social Facebook e classificar o sentimento desta em positivo, negativo ou neutro. Para tanto, o sistema possui as seguintes funcionalidades: coleta dos dados, pré-processamento dos dados e análise dos sentimentos.

**Coleta dos Dados:** Responsável por coletar mensagens da Fanpage selecionada pelo usuário no Facebook;

**Pré-Processamento:** Responsável por efetuar algumas correções no vocabulário e limpeza nas mensagens coletadas para se fazer uma melhor análise utilizando o Natural Language Toolkit (NLTK);

**Análise de Sentimentos:** Fase responsável por calcular o valor da polaridade de cada mensagem. O sistema exibe para o usuário a mensagem coletada, o valor sentimental (positivo, negativo ou neutro) e o valor da polaridade utilizando a API Senticnet;

Será demonstrado nas seções seguintes a vi-

são detalha do desenvolvimento de cada fase.

## Coleta dos Dados

Para o desenvolvimento da coleta de dados, utilizou-se a API Graph v2.9, principal API da plataforma. O Graph se baseia no protocolo HTTP e permite executar métodos (GET, POST ou DELETE) para consultar dados, publicar postagens, carregar fotos, dentre outras interações com o Facebook.

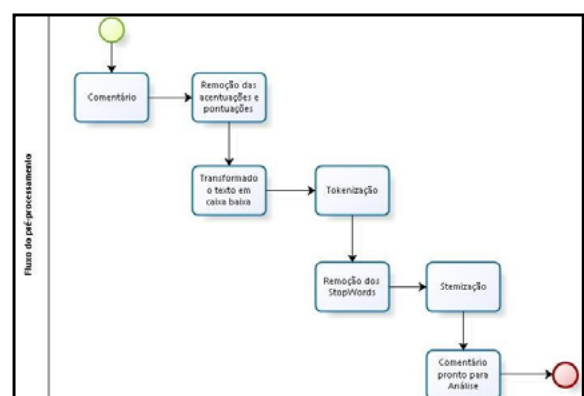
Com esta API disponibilizada pelo Facebook para desenvolvedores, é possível acessar qualquer recurso da rede social, para realizar o desenvolvimento dessa fase foi utilizado apenas o método GET, responsável por recuperar as informações necessárias para funcionamento da aplicação.

Ao final desta etapa, iniciou-se a etapa de pré-processamento, que é explicada na seção a seguir.

## Pré-processamento

Uma vez coletados, os textos são pré-processados para remoção de acentuação, pontuações, conversão para letras minúsculas (caixa baixa), divisão do texto em palavras (tokenização), redução de palavras flexionadas (stemming), palavras consideradas irrelevantes para o processo de análise do sentimento foram eliminadas do texto (stop words), a fim de otimizar o desempenho da tarefa de classificação, conforme ilustrado na Figura 1. Durante o pré-processamento do texto utilizou-se a biblioteca Natural Language Toolkit (NLTK) v3.2.4. O NLTK 2 é uma biblioteca em Python para manipulação de textos em linguagem natural.

**Figura 1** - Fluxo do pré-processamento



## Análise de Sentimentos

Depois de realizar o pré-processamento, as postagens passam pelo processo de classificação. Para o desenvolvimento desta etapa foi utilizado a API SenticNet v4.0. Esta API utiliza um dicionário com mais de 30.000 conceitos de senso comum para categorizar textos utilizando 3 (três) técnicas: Polaridade, Sentic Level e Semantic Level. No SentEdu utilizou-se a classificação por Polaridade, que consiste em quebrar a mensagem em palavras e realizar uma busca palavra por palavra na base do SenticNet afim de encontrar o valor polaridade de cada. Ao terminar a varredura nas palavras é feito um somatório dos valores das polaridades encontradas e calculado a média, aonde essa média indica a polaridade da mensagem como um todo.

Nesta técnica, as palavras classificadas com polaridade no intervalo de  $[1,0]$  são consideradas como negativas, enquanto as palavras rotuladas com intervalo  $[0,1]$  são positivas. A polaridade da frase é dada pela média das polaridades de todas as palavras que compõe.

Por fim, o SentEdu apresenta ao usuário o resultado da avaliação contendo a mensagem, o valor da polaridade calculada e a classificação de sentimento (positivo, negativo ou neutra).

## INTERFACE DA APLICAÇÃO

Ao iniciar a aplicação, no primeiro acesso, é solicitada a autenticação do usuário com as credenciais do Facebook, conforme ilustrado na Figura 2(a). Uma vez autenticado, a aplicação apresenta a tela de "Busca por Páginas" (Figura 2(b)), em que o usuário ao clicar em "buscar" iniciará o processo de busca das suas páginas no Facebook. As páginas que o usuário tem permissão são identificadas e mapeadas em botões que permitem selecionar de qual página será coletada as postagens e avaliado o sentimento, conforme exibido na tela de "Páginas Disponíveis" (Figura 2(c)).

Após selecionar a página, a aplicação exibe as postagens (comentários) existentes nela, (Figuras 3(a) e 3(c)). Assim, o usuário poderá visualizar e navegar nos comentários, bem como submetê-los à avaliação, conforme ilustrado na Figura 3(b). Para

cada comentário avaliado é exibido na tela de "Avaliação dos Comentários" a descrição deste, a polaridade do sentimento e o seu valor dentro de uma caixa com plano de fundo verde, branco ou vermelho, a fim de facilitar o reconhecimento visual e o mapeamento do sentimento do comentário em positivo, neutro ou negativo, respectivamente.

Figura 2 - Telas iniciais do SentEdu



Figura 3 - Telas da Análise de Sentimento do SentEdu



## MATERIAL E MÉTODOS

Foi executado um estudo de caso com os discentes e docentes do IFS para avaliar o desempenho da análise de sentimento da ferramenta proposta. Para tanto, utilizou-se o SentEdu para coletar e inferir o sentimento das postagens de discentes, docentes e da comunidade em geral na página do Facebook do IFS, realizadas no período entre 03 de Novembro de 2013 a 19 de Maio de 2017, gerando um montante de 25 (vinte e cinco) comentários.

Em seguida, foi criado um questionário on-line, no Google Forms, contendo os 25 comentários coletados para que discentes e docentes do IFS avaliassem os sentimentos das postagens. Assim, para cada comentário, foi solicitado que os entrevistados julgassem se a postagem representava um sentimento positivo, negativo ou neutro. O questionário de avaliação ficou disponível no período de 21 de junho até 06 de julho de 2017 e obteve 18 respostas.

Após o período de coleta da avaliação, os resultados foram sumarizados considerando a maior frequência como critério para definição do sentimento da postagem. Ou seja, as postagens que receberam mais votos positivos foram consideradas como sentimento positivo; as publicações que receberam mais votos negativos foram mapeadas em negativo, assim como, as mensagens com maior frequência de classificação neutra foi definido como sentimento neutro. Os resultados desta avaliação foram comparados com a análise de sentimento do SentEdu, a fim de mensurar o precision (P), recall (R), accuracy (A) e f-measure (F) para a aplicação desenvolvida. Durante a avaliação utilizou-se ainda a matriz de confusão para visualizar os Falsos-Positivos (FP), os Falsos-Negativos (FN), os Verdadeiros-Positivos (VP) e os Verdadeiros-Negativos (VN).

## RESULTADOS E DISCUSSÃO

Os resultados apresentados pelo sistema mostraram-se compatíveis com os apontados pela pesquisa. Na classificação feita pelo sistema, 19 comentários foram classificados como sendo positivo, 2 como negativos e 4 como neutros. Já na classificação feita pelas pessoas através da pesquisa, 18 comentários foram classificados como positivos, 4 como negativos e 3 como neutros. Ao realizar o comparativo dos resultados identificou-se que o sistema classificou corretamente 60% das postagens, ou seja, acertou 15 dos 25 comentários. Na Tabela 2 está apresentada a matriz de confusão da classificação dos comentários realizada através da pesquisa e do SentEdu.

**Tabela 2** - Telas iniciais do SentEdu.

		Humana		
		Positivo	Neutro	Negativo
SentEdu	Positivo	14	2	3
	Neutro	2	1	1
	Negativo	2	0	0

Analisando os resultados utilizando as métricas de desempenho de precision (P), recall (R), accuracy (A) e f-measure (F) constatou-se que o sistema atingiu maior precisão na classificação de comentários positivos, dos 18 comentários classificados, 14 deles foram corretamente categorizados, o que fez com que o desempenho de precisão atingisse 78%, conforme ilustrado na Tabela 4.

**Tabela 4** - Análise do SentEdu com as métricas de desempenho.

	Positivo	Neutro	Negativo
V.P	14	0	1
F.P	4	4	2
V.N (Neu.)	1	1	0
V.N (Neg.)	0	14	14
F.N	5	2	3
Accuracy (%)	60%	60%	60%
Recall (%)	74%	0%	25%
Precision (%)	78%	0%	33%
F-measure (%)	76%	0%	28%

O recall corroborou com este resultado da precisão e destacou que os melhores resultados foram na avaliação dos comentários positivos. O sistema possui uma accuracy de 60%, sendo este o total de acerto em relação a todas as categorias, o mesmo já mencionado anteriormente. Por fim, foi realizada a medida geral de eficiência do desempenho do sistema levando em consideração os erros de classificação em todas as três catego-



rias de polaridade. Por meio do cálculo da F-Measure percebeu-se que os melhores resultados foram obtidos na classificação dos comentários positivos, que obteve um resultado de 76%.

Nota-se, portanto, que os resultados foram bem próximos em relação aos sentimentos classificados como positivo, o que mostra um ponto positivo para o sistema, enquanto para as demais classificações não foram obtidos resultados satisfatórios. Na Tabela 5 é possível entender os problemas enfrentados durante a classificação.

**Tabela 5** - Análise comparativa das postagens avaliadas

Comentário	Sistema	Pesquisa
“Qualidade e excelência! Foi muito bom ter conhecido uma equipe tão competente comprometida com a Educação! Bravo.”	Positivo	Positivo
“Muito bom. Ótimos professores! E a estrutura física promete melhorar ainda mais.”	Negativo	Positivo
“To muito decepcionada por não poder estudar devido a falta de alunos.”	Positivo	Negativo
“Profissionais qualificados e preparados para o ensino tecnológico, uma equipe de Excelência, PARABENS!”	Neutro	Positivo

No primeiro comentário, apresentado na Tabela 5, há um resultado de classificação exata pelo sistema, enquanto o segundo comentário apresenta valores divergentes na classificação, os termos "estrutura" e "melhorar" possuem respectivamente uma polaridade positiva de 0.089 e 0.243. Já o termo "ainda" possui uma polaridade negativa de -0.522, valor que define o sentimento negativo do comentário.

Para a análise humana, a polaridade desse comentário foi classificada como positiva, entretanto, a mensagem possui o termo "ainda" que predomina no valor da classificação de toda a expressão. O terceiro comentário apresenta a mesma justificativa. Já o quarto comentário, foi classificado pelo sistema como neutro, visto que a aplicação não conseguiu identificar um resultado para o mesmo. Através desse exemplo, é possível perceber as dificuldades encontradas no processo de análise de sentimento, considerando que o comentário possui uma polaridade positiva segundo a pesquisa.

## CONCLUSÕES

Este trabalho apresentou o SentEdu, uma aplicação móvel para coletar mensagens da página do Facebook do IFS, análise de sentimento baseada em polaridade com a API SenticNet. O objetivo da ferramenta é permitir aos gestores do IFS acompanhar e analisar, periodicamente, o sentimento da comunidade sobre a imagem da Instituição, a fim de sustentar futuras tomadas de decisão.

Para avaliar a eficácia da aplicação, os comentários da Fanpage da Instituição foram coletados, avaliados e comparados com a análise de docentes e discentes do próprio Instituto. Os resultados obtidos foram satisfatórios, pois a aplicação classificou corretamente 60% dos comentários, embora a base de vocabulário da língua portuguesa do SenticNet seja bastante limitada e o quantitativo de avaliadores humano tenha sido pequeno.

Porém, mesmo com o resultado alcançado, identificou-se limitações na ferramenta que podem ser melhoradas como, por exemplo, erros de digitação e palavras que não constavam no "vo-

cabulário" da API SenticNet, fazendo com que comentários fossem classificados erroneamente como neutros.

## REFERÊNCIAS

LIU. **Web data mining: exploring hyperlinks, contents, and usage data.** Springer Science & Business Media. 2007.

Malheiros, Y., Lima, G., and Tinto-PB-Brasil, R. **Uma ferramenta para análise de sentimentos em redes sociais utilizando o senticnet.** 2013.

Moreira, V. d. S., Siqueira, S. W., Andrade, L., and Pimentel. **Análise de sentimentos: Comparando o uso de ferramentas e a análise humana.** 2016.

Pang, B. and Lee, L. Opinion mining and sentiment analysis (foundations and trends (r) in information retrieval). 2008.

Tedeschi, A. and Benedetto, F. A cloud-based big data sentiment analysis application for enterprises' brand monitoring in social media streams. In Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), 2015 IEEE 1st International Forum on, pages 186–191. IEEE.

Rodrigues, C. A. S. **Mineração de Opinião / Análise de Sentimentos.** 2010.

Santos, L. M. **Protótipo para mineração de opinião em redes sociais: Estudo de casos selecionados usando o twitter.** 2010.