Metodologia TCC

Realizar análises automáticas de frases através de um algoritmo requer métodos específicos e por essa razão nos apoiamos em outros projetos que seguiram neste mesmo intuito da análise de sentimentos desenvolvendo um instrumento técnicas de análise de sentimentos em textos não estruturados, detectando o estado de humor de frases escritas por pessoas.

 A tese “Análise de dados de textos curtos provenientes de redes sociais” (SILVA, 2016) desenvolveu métodos para validar um algoritmo que aplica as técnicas de análise de sentimentos automáticas em *tweets*. Este projeto realizou experimentos com fundamentos voltados para texto de *tweets,* mas com uma análise realizada por nós percebemos que poderíamos adaptar alguns conceitos das técnicas utilizadas pela autora para o nosso projeto. As técnicas extraídas desta tese foram as que eram baseadas em léxicos e paradigmas supervisionados. “Métodos baseados em léxicos obrigatoriamente necessitam de uma lista de palavras previamente computadas"(SILVA, 2016). Ainda segundo SILVA (2016) “Para métodos de análise de sentimentos definidos sob o paradigma supervisionado, é imprescindível a definição prévia de um conjunto de treinamento, em que um especialista de domínio atribui a cada tweet analisado um rótulo de classe de sentimento”.

* Exemplos de aplicações

As técnicas de análise de sentimentos abrangem áreas de bastante relevância, como por exemplo o setor financeiro. O mercado financeiro tem como um de seus conceitos o gerenciamento de risco, em específico o mercado de títulos. Automatizar o processo de análise de mercado, através de análise de texto, ajuda a antecipar a direção dos ativos, bem como o nível de incerteza (volatilidade) antes de realizar um determinado investimento (PEZZINI, 2017). O autor PEZZINI (2017) desenvolveu projeto nesse setor, aprimorando dicionários léxicos para classificar textos específicos sobre o mercado financeiro e verificando o limite de classificação que o léxico pode alcançar referente a textos já classificados.

Verificados então, que realizando as devidas adequações e ajustes à nossa proposta, podíamos aplicar, no desenvolvimento do software, as técnicas de análise de sentimentos tomando como base léxicos e a análise de rótulos supervisionados.

 A metodologia que aplicamos iniciou com o levantamento do tema para ser aplicado na técnicas de análise de sentimentos a um determinado tema, com o objetivo de testar um software usando um tema específico. A busca pelo tema teve como critério qual campo de âmbito da comunidade acadêmica poderia ser contemplado com um suporte tecnológico que envolve a análise de sentimentos. Levantamos também dados que comprovem a nossa abordagem da problemática em que afirmam que a evasão de alunos de cursos de níveis superiores é de fato preocupante.

Em seguida pesquisamos, decidimos e adotamos um tema bastante interessante que é um problema recorrente nas instituições de ensino que é a evasão de alunos dos cursos de nível superior. Com o tema escolhido, analisamos outro tópico para decidirmos qual curso e disciplinas seriam aplicados ao nosso projeto. Os critérios para escolha do curso e das disciplinas foram: cursos de tecnologia e disciplinas com maior evasão conforme a problemática abordada neste projeto, a análise dos dados abordados do CBSI (2018). Então com base nessas análises o curso escolhido foi o TADS (Tecnologia em Análise e Desenvolvimento de Sistemas) e as disciplinas escolhidas foram: Programação I, TCC 2, Libras e Engenharia de software 1.

Com a instituição de ensino, curso, disciplinas e o tema escolhido fizemos então os levantamentos bibliográficos em projetos que teve como proposta analisar os sentimentos de alunos que estejam estudando e enfrentando dificuldades  acadêmicas e que esses projetos contribuíssem de alguma forma para auxílio no combate a evasão de alunos das instituições de ensino. E o projeto mais próximo da nossa proposta e relevante aos critérios do nosso levantamento bibliográfico, foi o projeto do autor YU, L. C. et al. com o projeto *Improving early prediction of academic failure using sentiment analysis on self-evaluated comments*. Publicado em: **Journal of Computer Assisted Learning**, v. 34, n. 4, p. 358–365, ago. 2018.

Esse projeto contribuiu para nortear nossas pesquisas em relação a coleta de dados e ao software implementado. A estrutura do formulário foi inspirada no instrumento que o autor elaborou para coletar os dados para processar e chegar ao resultado esperado no seu projeto. Então construímos o formulário adaptado para nossa realidade que é constituído de dois campos em que o aluno preenche com as suas informações de sentimentos. O primeiro campo é a escolha da rotulagem de uma frase que baseada em polaridade de sentimento. Esse rótulo representado por três figuras, e cada figura representa um tipo de estado de humor específico, que são o negativo, neutro e positivo (cf. figura 1).



Figura 1

Essa rotulagem ajuda os alunos a identificar e marcar a figura segundo o sentimento da frase escrita no segundo campo. Esse segundo campo é onde o aluno expressa uma breve autoavaliação sobre seu desempenho pós-avaliativo (cf. figura 2).

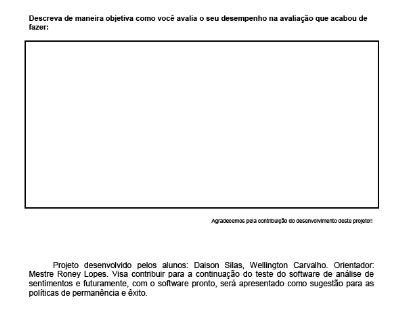


Figura 2

Após desenvolvido o formulário iniciamos a aplicação nas disciplinas de TADS, as quais eram autorizados pelo professores para que fosse aplicado. Nas disciplinas em que o formulário era entregue também era entregue um termo em que o aluno dá permissão para usarmos os dados para fins de pesquisa.

Outras ferramentas de coleta analisadas, como por exemplo, formulários enviados via e-mail ou pela Web, foi considerado desinteressante para o projeto pelo fato de que correríamos o risco de não ter a compressão e participação de grande parte dos estudantes.

Durante a aplicação em sala de aula foi realizada a apresentação da proposta do projeto e do formulário para alunos das disciplinas do curso de TADS. Explicamos como preencher as informações indicando de forma objetiva a sua percepção (Positivo, Negativo ou Neutro). Os formulários preenchidos ou não, até porque o aluno tinha a opção de responder ou não, foram todos coletados e encaminhados para a etapa de digitalização. Entende-se que ao obter essa auto-avaliação durante a realização do instrumento avaliativo, é possível captar o estado de humor do estudante em seu estágio de maior exposição.

 A digitalização dos formulários foi de forma manual, digitamos formulário por formulário para arquivos de texto utilizando o editor “*Sublime Text 3*” que posteriormente foram salvos em arquivos do tipo *“.TXT”*. Após este processo de digitalização de todos os formulários foi gerado um banco de dados, sendo submetido ao artefato de software que realizaria a análise de sentimentos.

Para o uso das técnicas de análise de sentimentos foi necessário seguir processos de desenvolvimento de software, uso de bibliotecas e entender como é granulada uma mensagem. Para a primeira etapa do software foi necessário iniciar com alguns protótipos que nos permitiu realizar testes e analisar o que precisaria ser melhorado para a atender a proposta do projeto. Os requisitos de maior relevância para o software tais como bibliotecas, léxico e plataforma (IDE) para desenvolvimento e conclusão do software foram analisados, selecionados e utilizados. A metodologia do software foi da seguinte maneira: