

Análise de dados e de sentimentos para melhoria de resultados em Sistemas de Avaliação Institucional

**Gustavo Kataoka¹, Amaury Antônio de Castro Junior¹,
Anderson Corrêa de Lima¹, Caroline Pauletto Spanhol²**

¹Faculdade de Computação (Facom) – Universidade Federal de Mato Grosso do Sul (UFMS)
Campo Grande – MS – Brazil

²Escola de Administração e Negócios (ESAN) – Universidade Federal de Mato Grosso do Sul
(UFMS) Campo Grande – MS – Brazil

gustavo.kataoka@ufms.br, amaury.junior@ufms.br,
anderson.lima@ufms.br, caroline.spanhol@ufms.br

Abstract. *This work aims to develop improvements in the Institutional Assessment System of the Federal University of Mato Grosso do Sul. To this end, a survey was carried out among its users with the aim of discovering the needs and opportunities for improvement to be implemented in the system. The survey revealed two needs, the first being the automation of a report that is carried out using a spreadsheet, and the second the implementation of a classifier of textual responses provided by system respondents. In addition to the meeting with users, a systematic literature review was carried out to investigate whether the desired improvements were common to other institutions and what techniques and methods were used to resolve them. The bibliographical research revealed that, in textual analysis, some institutions use artificial intelligence and machine learning techniques. This work, in turn, presents, as an improvement to the system, the development of the Action Report and the sentiment analyzer using BERT. A Time Series report discovered in RSL was included as an improvement developed in this work. With the implementation of these improvements in the system, it is expected to make the work routine of its users more efficient and productive, helping them in decision making.*

Resumo. *Este trabalho tem como objetivo desenvolver melhorias no Sistema de Avaliação Institucional da Universidade Federal de Mato Grosso do Sul. Para isso, foi realizado um levantamento junto aos seus utilizadores com o intuito de descobrir as necessidades e oportunidades de melhoria a serem implementadas no sistema. O levantamento revelou duas necessidades, sendo a primeira a automação de um relatório que é realizado com uso de planilha, e a segunda a implementação de um classificador de respostas textuais fornecidas por respondentes do sistema. Além da reunião com os utilizadores, uma revisão sistemática da literatura foi realizada para investigar se as melhorias desejadas são comuns a outras instituições e quais foram as técnicas e métodos empregados na sua resolução. A pesquisa bibliográfica revelou que, nas análises textuais, algumas instituições utilizam inteligência artificial e técnicas de aprendizagem de máquina. Este trabalho, por sua vez, apresenta, como melhoria para o sistema, o desenvolvimento do Relatório de Ações e o analisador de sentimentos utilizando o BERT. Um relatório de Séries Temporais descoberto na RSL foi incluído como melhoria desenvolvida neste trabalho. Com a implementação destas melhorias no sistema espera-se tornar a rotina de trabalho dos seus utilizadores mais eficiente e produtiva, auxiliando-os na tomada de decisão.*

1. INTRODUÇÃO

A avaliação, em geral, é concebida como um meio de diagnosticar problemas, a fim de promover mudanças visando a melhoria da qualidade dos serviços prestados à comunidade, constituindo um mecanismo para sua melhoria contínua, além de um instrumento de prestação de contas à sociedade [Paula 2004].

A avaliação institucional das Instituições de Ensino Superior do Brasil (IES) é regulamentada pelo Sistema Nacional de Avaliação da Educação Superior (SINAES), pela Lei nº 10.861, de 14 de Abril de 2004¹, que tem por finalidade a melhoria da qualidade da educação superior, a orientação da expansão da sua oferta, o aumento permanente da sua eficácia institucional e efetividade acadêmica e social e, especialmente, a promoção do aprofundamento dos compromissos e responsabilidades sociais das instituições de educação superior, por meio da valorização de sua missão pública, da promoção dos valores democráticos, do respeito à diferença e à diversidade, da afirmação da autonomia e da identidade institucional (Nunes et al., 2017).

A avaliação institucional, regulamentada pelo SINAES, consiste em avaliação externa e avaliação interna, ou simplesmente, autoavaliação. A autoavaliação é realizada por meio da consulta aos segmentos da comunidade acadêmica com a utilização de questionários que buscam obter o *feedback* de diversos aspectos do serviço prestado pela instituição. Pode-se entender como segmentos as pessoas que compõem a comunidade acadêmica, sendo alunos, técnicos administrativos, professores, coordenadores de curso e diretores (Sousa et al., 2022).

Na Fundação Universidade Federal de Mato Grosso do Sul - UFMS² a realização da autoavaliação é feita por meio do Sistema de Avaliação Institucional (SIAI)³, um sistema informatizado projetado e mantido pela Agência de Tecnologia da Informação e Comunicação - Agetic/UFMS⁴ para aplicação dos questionários aos segmentos da comunidade acadêmica.

O SIAI tem sido amplamente utilizado na instituição, mas carece de algumas ferramentas capazes de auxiliar o trabalho da Comissão Própria de Autoavaliação (CPA), responsável pelo processo de autoavaliação na UFMS. Nesse sentido, o objetivo deste trabalho foi desenvolver melhorias no SIAI que contribuíssem na elaboração de relatórios e na tomada de decisão. Para atingir este objetivo geral, foram executados os seguintes objetivos específicos:

1. Identificar as principais necessidades de melhorias nos relatórios do SIAI junto aos seus utilizadores;
2. Verificar, por meio de uma Revisão Sistemática da Literatura (RSL), se as necessidades levantadas são comuns em outras organizações, observando as técnicas e metodologias empregadas na resolução do problema;
3. Relacionar as técnicas e métodos identificados na pesquisa que podem ser aplicáveis ao SIAI;

¹ https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/110.861.htm

² <https://www.ufms.br/>

³ <https://siai.ufms.br/avaliacao-institucional/>

⁴ <https://agetic.ufms.br/>

4. Desenvolver as necessidades identificadas utilizando a RSL como ponto de partida.

1.1. Justificativa

A justificativa para este trabalho é tornar o trabalho da CPA, principal gestora e interessada do sistema, mais eficiente ao fazer com que o SIAI consiga produzir informações e relatórios, que hoje são produzidos com auxílio de aplicativos externos a ele. Isso gerará a otimização do tempo e evitará desperdícios com recursos humanos na confecção desses relatórios, além de evitar erros humanos em sua produção e auxiliar na tomada de decisão.

1.2. Metodologia

Para atingir o objetivo deste trabalho em desenvolver melhorias no SIAI, foi necessário conhecer quais são as melhorias desejadas pelos utilizadores deste sistema.

Após o levantamento das melhorias desejadas, uma RSL foi utilizada como ponto de partida na busca por soluções e métodos empregados por outras organizações na resolução destas necessidades, no caso de serem comuns.

Identificadas as formas de como resolver o problema, o próximo passo foi iniciar o desenvolvimento das melhorias, e por fim realizar a entrega do sistema melhorado.

2. SIAI

O SIAI é um sistema desenvolvido pelo setor de desenvolvimento de software da UFMS. Seu propósito visa o atendimento da autoavaliação da instituição, atuando como um sistema de questionário que é aplicado à comunidade acadêmica.

Ao acessar o sistema como respondente, este apresenta questões estruturadas com escala de resposta do tipo *Likert*⁵. Também existem questões abertas, nas quais os respondentes têm a liberdade de expressarem suas opiniões e sentimentos livremente.

O sistema foi desenvolvido como um aplicativo Web, utilizando o conceito de separação em *backend* e *frontend*. Neste tipo de desenvolvimento, o *frontend* refere-se a uma aplicação criada em uma linguagem interpretada pelo navegador Web (geralmente Javascript, Html e CSS) encarregada da interação direta com o usuário, tanto para apresentação quanto manipulação das informações do sistema. Enquanto que o *backend* trata-se de uma aplicação que roda em um servidor e fica encarregado de atender as solicitações vindas do *frontend*, abrigando a maioria das regras de negócio do sistema, incluindo a manipulação do banco de dados.

O *frontend* do SIAI foi criado utilizando o VueJs⁶ 2.x., um *framework open source* em Javascript para desenvolvimento de interfaces de usuário. Já o *backend* foi desenvolvido utilizando o Spring Boot⁷ 3.2.x, um *framework* em linguagem de programação Java para criação de microsserviços.

⁵ Técnica de pesquisa para medir opiniões e sentimentos em relação a um determinado tema.

⁶ <https://vuejs.org/>

⁷ <https://spring.io/projects/spring-boot>

2.1. Melhorias solicitadas

As melhorias solicitadas são duas: A automatização do relatório de ações e a classificação das respostas textuais fornecidas pelos respondentes. Essas necessidades foram levantadas por meio de reuniões realizadas com a CPA e são descritas nas subseções 2.1.1 e 2.1.2.

2.1.1. Relatório de ações

A CPA, em sua atuação como gestora do SIAI, relatou sobre como é complicada a rotina após a coleta de dados com o sistema junto aos respondentes. Nessa fase, o foco das operações com o sistema está na extração das informações coletadas durante o período de aplicação dos questionários, a fim de compilar um relatório contendo as ações a serem realizadas que colocarão a UFMS em rumo a melhoria da qualidade de seus serviços prestados à comunidade.

Os dados de avaliação podem e devem ser utilizados para fins de melhoria, sendo utilizado não somente como um instrumento regulatório, mas também como um instrumento que aponta onde a instituição está acertando e, principalmente, onde está errando para corrigir esses fatores apontados.

Nessa fase, o SIAI possui os dados necessários para fazer esses apontamentos, porém seus relatórios fornecem aos seus utilizadores apenas dados brutos, ou seja, apresenta os dados na forma em que foram coletados, não fornecendo nenhum tipo de análise sobre esses dados.

Desse modo, a CPA, juntamente com as Comissões Setoriais de Autoavaliação (CSA), realizam uma extração dos dados do SIAI e, utilizando um software de processamento de planilhas (geralmente o Excel), produzem um relatório externo ao sistema, copiando esses dados para dentro da planilha. Então, a planilha realiza alguns cálculos sobre os dados, tais como o cálculo da média e mediana.

Ao ter posse dos dados, a planilha é utilizada para confeccionar um relatório de ações para cada questão apresentada na avaliação e agrupada por seu eixo. Esse relatório segue a premissa de que ações devem ser observadas para medianas abaixo de 4, possuindo, ainda, uma classificação para a ação, sendo que para mediana igual a 3 considera-se como “Razoavelmente preocupante”, e abaixo disso como “Muito preocupante”.

O relatório apresentado é então utilizado para obter os apontamentos sobre quais ações devem ser tomadas para direcionar a instituição na melhoria de seu desempenho. Os utilizadores do SIAI desejam que esse relatório seja emitido diretamente pelo sistema, sem precisar de softwares auxiliares para confecção do mesmo. O anseio da CPA pela demanda deste relatório pode ser considerado como o originador deste trabalho.

2.1.2. Análise de sentimentos em questões abertas

Outra demanda apontada pela CPA é direcionada para as respostas às questões descritivas, também referenciadas como questões abertas, nas quais os respondentes de

cada segmento têm a possibilidade de relatar sobre o item de forma textual.

As respostas descritivas devem ser analisadas pela CPA por meio da leitura das mesmas. Dependendo do número de questões descritivas aplicadas, o volume de respostas a serem analisadas pode ser alto. Apesar desse trabalho ser dividido entre as comissões setoriais da UFMS, analisar todas as respostas, em sua completude, é muito trabalhoso e complexo, demandando muito tempo dos colaboradores envolvidos.

Para facilitar o trabalho de seus utilizadores, a CPA propõe munir o SIAI com um classificador de textos responsável por definir automaticamente as respostas das questões descritivas segundo os seguintes rótulos: Denúncia, Elogio, Sugestão, Reclamação.

3. REVISÃO SISTEMÁTICA DA LITERATURA

A RSL foi realizada por meio da utilização do software Parsifal⁸, que é baseado no guia de procedimentos para realização de RSL em Engenharia de Software proposto por Kitchenham, B. and Charters, S. (2007). Ou seja, a RSL foi realizada por meio da utilização de palavras-chave introduzidas pelo acrônimo PICOC, do inglês *Population, Intervention, Comparison, Outcome and Context*, ou População, Intervenção, Comparação, Resultado e Contexto, e posteriormente foram definidas as estratégias de filtragem e seleção de trabalhos relacionados ao tema.

A estratégia de busca é baseada na utilização dos termos utilizados na População, Intervenção e Resultado em conjunto com sinônimos em língua inglesa. Após, foi realizada a junção desses termos utilizando operadores lógicos *AND* e *OR*, resultando na cadeia de busca apresentada a seguir.

**(avaliação AND institucional) OR (avaliação AND
institucional AND melhoria) OR (institucional AND
assessment AND improvement) OR (institucional AND
evaluation AND improvement)**

Para realização da pesquisa foram utilizadas as bibliotecas digitais apresentadas na Tabela 1. A pesquisa nas fontes relacionadas revelou a base primária apresentada na coluna “Encontrados”.

Tabela 1. Estudos encontrados nas fontes de pesquisa.

Fontes	URL	Encontrados	Selecionados
ACM Digital Library	http://dl.acm.org/	50	4
Google Scholar	http://scholar.google.com/	70	37
IEEE Explore	http://ieeexplore.ieee.org	68	4
Periódicos da Capes	https://www.gov.br/capes/pt-br	115	29
Scientific Electronic Library Online	https://www.scielo.br/	218	40
Total		588	114

⁸ Parsifal (<https://parsif.al/>): Ferramenta online desenvolvida para apoiar pesquisadores na realização de revisões sistemáticas de literatura.

Após a realização da busca, foi realizado o refinamento da pesquisa em duas etapas. A primeira tendo como base o acesso ao artigo e leitura do título, seu(s) autor(es) e resumos. Os critérios para inclusão e exclusão na primeira etapa partiu da checagem apresentada na Tabela 2, sendo o resultado desta etapa apresentado na coluna “Selecionados” da Tabela 1.

Tabela 2. Critérios de inclusão e exclusão.

Inclusão	Exclusão
Título fazer inferências a avaliação institucional	Duplicações encontradas
Utilização de artigos revisados por pares	Título cita avaliação, mas fora do contexto
O arquivo estar acessível para leitura	Utilização de idioma diferente de português ou inglês

A segunda etapa do refinamento da pesquisa foi embasada em buscar pelos trabalhos que pudessem responder aos questionamentos listados abaixo. Para cada questão que o trabalho tivesse a chance de conter a resposta foi atribuído um peso igual a 1.0. Caso respondesse de forma parcial, o peso atribuído foi de 0.5. E para os casos em que não fornecesse algo que contribuísse para a questão, foi atribuído peso igual a 0 (zero).

- O artigo cita sistema informatizado para avaliação?
- O artigo aborda melhorias utilizando o instrumento de avaliação?
- O artigo aborda Avaliação institucional?
- Usa tecnologia própria e recente? (Foi considerado como recente uma tecnologia de até 5 anos)
- Aborda métodos de extração e como é utilização dos dados de avaliação?

Para realizar a distribuição de pesos em cada questão nos trabalhos selecionados, foi feita uma leitura superficial em cada um deles. A soma de todos esses pesos para cada trabalho foi considerada como uma nota geral para esse trabalho em que, quanto mais próxima de 5 (maior nota a ser atingida), mais o trabalho poderia contribuir com este estudo.

Uma nota de corte foi utilizada como critério de seleção. A nota de corte definida foi 2.5, sendo escolhida por meio da observação dos trabalhos encontrados, visto que seria esta que identificaria ao menos duas das respostas como afirmativas acrescida de uma parcial ou cinco como parciais, sendo considerado como o mínimo necessário para que o trabalho tivesse algo a corroborar com esta pesquisa segundo a visão do autor.

A partir da nota de corte, um total de 19 artigos foram selecionados e utilizados para realizar a revisão da literatura.

A Tabela 3 mostra a tabulação dos dados da RSL. Essa tabela apresenta os desafios no tema de Avaliação Institucional e as possíveis soluções apontadas pelos autores pesquisados.

Tabela 3: Desafios e soluções na área de Avaliação Institucional encontrados na RSL

Demandas/Desafios	Proposta/Solução	Autores
Classificação das respostas	Analisador de sentimentos utilizando Naive Bayes, K-Nearest Neighbor, Árvore de Decisão, Support Vector Machine (SVM), K-means	<ul style="list-style-type: none"> • Sousa et. al. (2019)
	Analisador de sentimentos utilizando redes neurais artificiais	<ul style="list-style-type: none"> • Barbosa et.al. (2019)
Disponibilidade dos dados	Aumento de acessibilidade do sistema de informação	<ul style="list-style-type: none"> • Luz et.al. (2020)
Sensibilização dos segmentos	Envio de email convidando a responder a avaliação	<ul style="list-style-type: none"> • Paula et.al. (2004) • Reis et.al. (2010) • Nascimento et.al. (2016) • Arruda et.al. (2019) • Silva et.al. (2019)
Confiabilidade nos dados	Alpha de Cronbach: técnica utilizada para avaliação da confiabilidade e consistência interna de instrumentos de medição	<ul style="list-style-type: none"> • Paula et.al. (2004) • Reis et.al. (2010) • Nascimento et.al. (2016) • Arruda et.al. (2019)
	Teoria da resposta ao item: relação entre a probabilidade de o participante responder corretamente a uma questão, seu conhecimento na área em que está sendo avaliado e as características dos itens.	
Descontinuidade de séries históricas	Análise diacrônica: análise baseada em evolução histórica das variáveis e parâmetros envolvidos em uma pesquisa e pode ser utilizada na resolução da descontinuidade de séries históricas.	<ul style="list-style-type: none"> • Kozen et.al. (2016) • Monticelli et.al.(2021)
Análise de dados	Análise fatorial exploratória (AFE): um conjunto de técnicas estatísticas amplamente utilizadas nas pesquisas em Psicologia. Durante sua execução, diversas decisões precisam ser tomadas a fim de se obter uma estrutura fatorial adequada.	<ul style="list-style-type: none"> • Nascimento et.al. (2016) • Jesus (2018)
	Análise fatorial confirmatória (AFC): esta técnica permite a verificação de ajustes entre os dados observados e um modelo hipotetizado a priori, o qual é baseado na teoria que especifica as relações causais hipotéticas entre fatores latentes (variáveis não observáveis) e suas variáveis indicadoras (observáveis).	
	Análise de variância: uma técnica estatística usada para comparar as variâncias entre as médias de diferentes grupos.	<ul style="list-style-type: none"> • Reis et.al. (2010)
	Análise de cluster: permite agrupar os tópicos conforme avaliação realizada sobre a Instituição.	
	Uso de Indicadores: definição dos parâmetros que servirão de base para a comparação entre a realidade de um objeto e seu estado desejável. A estes parâmetros são aferidos valores ligados aos seus objetivos, conferindo ao objeto avaliado um sentido prático. Então, é realizada uma comparação desses dados com o resultado esperado, identificando as variações em relação aos parâmetros. Após, procura-se detectar as causas dessas variações a fim de realizar medidas corretivas para promoção da melhoria do objeto avaliado.	<ul style="list-style-type: none"> • Shriberg (2002) • Erasmo et.al. (2018) • Nunes et.al. (2017) • Barbisan (2002)

A RSL foi realizada com o objetivo de descobrir se outras instituições possuem os mesmos problemas levantados na etapa de coleta de requisitos deste trabalho, e como estes problemas foram solucionados pelos autores pesquisados.

O resultado obtido com a RSL contribuiu para revelar técnicas e métodos que podem ser aplicados na problemática da análise de sentimentos de questões abertas, e também revelou novas perspectivas de trabalho a serem implementados no SIAI.

Dentre as análises utilizadas pelos autores pesquisados, destacou-se a **Análise Diacrônica**. Uma análise de dados baseada em séries temporais. Um relatório sobre os dados da avaliação que indica se houve melhoria ou não no decorrer do tempo será muito útil para auxiliar a tomada de decisões por parte dos utilizadores do sistema.

Esta análise foi incluída como requisito a ser implementado neste trabalho, por se tratar de um relatório essencial sobre os dados de avaliação institucional. A RSL apontou que os dados de avaliação devem ser utilizados para promover a melhoria do desempenho da instituição, e a Análise Diacrônica propõe atender esta demanda.

Em relação ao requisito de análise de sentimentos das questões descritivas, a RSL identificou que os trabalhos de Sousa et. al. (2019) e Barbosa et.al. (2019) abordaram este tipo de análise.

Segundo Barbosa et.al.(2019), o modelo que obteve melhor resultado em seu trabalho foi o BERT (Bidirectional Encoder Representations from Transformers), ao realizar a análise de sentimento de avaliação de cursos de graduação. Trata-se de um modelo baseado em redes neurais artificiais profundas e derivado dos *Transformers*.

Segundo Sun et.al. (2019), o modelo BERT tem alcançado excelentes resultados na resolução de diversos problemas na área de Processamento de Linguagem Natural (*Natural Language Processing* - NLP), incluindo a análise de sentimentos.

Portanto, para a resolução do problema de classificação das respostas descritivas, neste trabalho optou-se pela utilização do BERT pré-treinado em língua portuguesa criado por Souza et.al. (2020) e apelidado de BERTimbau⁹ por seu criador. Este modelo encontra-se disponível na plataforma *Hugging Face*¹⁰.

A adoção do BERTimbau foi vista como uma vantagem, já que o modelo foi pré-treinado em língua portuguesa e as respostas descritivas do SIAI também são coletadas neste idioma. Mais detalhes sobre esta solução serão abordados na seção de desenvolvimento dos requisitos deste trabalho.

4. DESENVOLVIMENTO DOS REQUISITOS

Esta seção apresenta os requisitos desenvolvidos para o SIAI, isto é, o Relatório de ações; os Relatórios de séries temporais; e a Análise de sentimentos de respostas descritivas. As subseções seguintes apresentam quais as estratégias e metodologias utilizadas para o desenvolvimento das melhorias no sistema, bem como os resultados obtidos na implementação.

4.1. Relatório de ações

O relatório de ações necessita do cálculo do valor mediano obtido na avaliação. A partir desse resultado, a planilha seleciona os itens avaliados que obtiveram resultado inferior ou igual a 3 (satisfatório, segundo a escala utilizada) para compor o relatório.

Para atingir o resultado desejado diretamente por meio do sistema, as seguintes tarefas foram realizadas no desenvolvimento:

- Calcular a mediana
- Criação de um filtro no relatório de questões objetivas

⁹ <https://huggingface.co/neuralmind/bert-base-portuguese-cased>

¹⁰ <https://huggingface.co/> - Startup que tem como missão a democratização e compartilhamento de código aberto com enfoque principal na área de Inteligência Artificial.

- Emissão do relatório a partir do filtro selecionado

O relatório de questões objetivas já existente no sistema foi utilizado para realização da automação do relatório de ações, ou seja, não houve necessidade de criação de uma ação específica no sistema para consulta dessas informações visto que o sistema já as fornece.

Ao passo que o sistema já calcula a mediana dos itens avaliados, o próximo passo foi a criação de um campo de filtro do relatório, no qual o usuário terá a possibilidade de informar qual a mediana de corte desejada para a confecção do mesmo.

A planilha utilizada na geração do relatório de ações de forma manual não dispõe de um campo para informar qual é a mediana de corte. Desse modo, a inclusão desse parâmetro no filtro de busca do relatório de questões objetivas impactará em uma maior flexibilidade para o usuário do sistema, de maneira que será permitido que seja definido o valor que se ajuste melhor a sua necessidade.

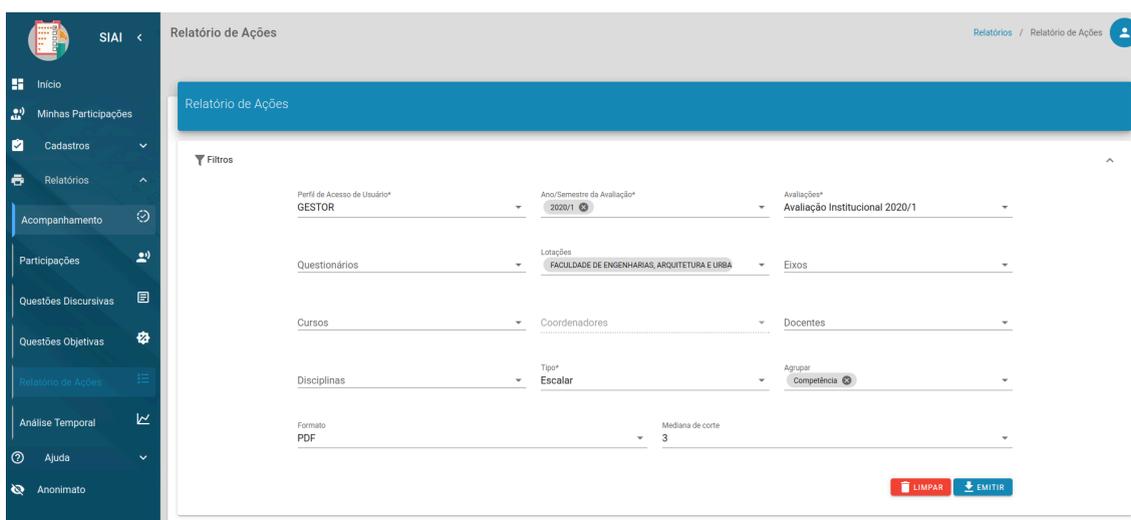
Para emitir o relatório, foi utilizado o serviço de consulta aos dados de questões objetivas. O usuário tem a opção de emitir um documento dentre os formatos disponíveis: XLS e PDF.

Como o backend da aplicação utiliza *Spring Boot* com Java, foi utilizado o *JasperReport*¹¹ para geração do documento.

Dois novos filtros foram, então, implementados no relatório, sendo: Agrupar e Mediana de corte. Dois filtros não são fáceis de manipular pela planilha que os utilizadores utilizavam anteriormente. Sendo possível agrupar por competência, lotação, curso e/ou disciplina.

Ao clicar em emitir na tela de filtros, o sistema devolve o relatório para o usuário no formato desejado segundo a opção selecionada, como mostra a Figura 1.

Figura 1: Tela de filtros para o relatório de ações do SIAI.



Fonte: Autor

¹¹ <https://community.jaspersoft.com/project/jasperreports-library> - Biblioteca Java para criação de documentos em diversos formatos.

4.2. Relatório de séries temporais

Como observado na RSL, a avaliação institucional é um importante instrumento para guiar a instituição na busca por melhorias em seus serviços prestados. Porém, este instrumento somente servirá a este propósito se seus dados forem apresentados para este fim. Ou seja, de nada adianta o sistema possuir essa informação, se seu utilizador não conseguir e/ou poder visualizá-la de forma a abstrair essa interpretação.

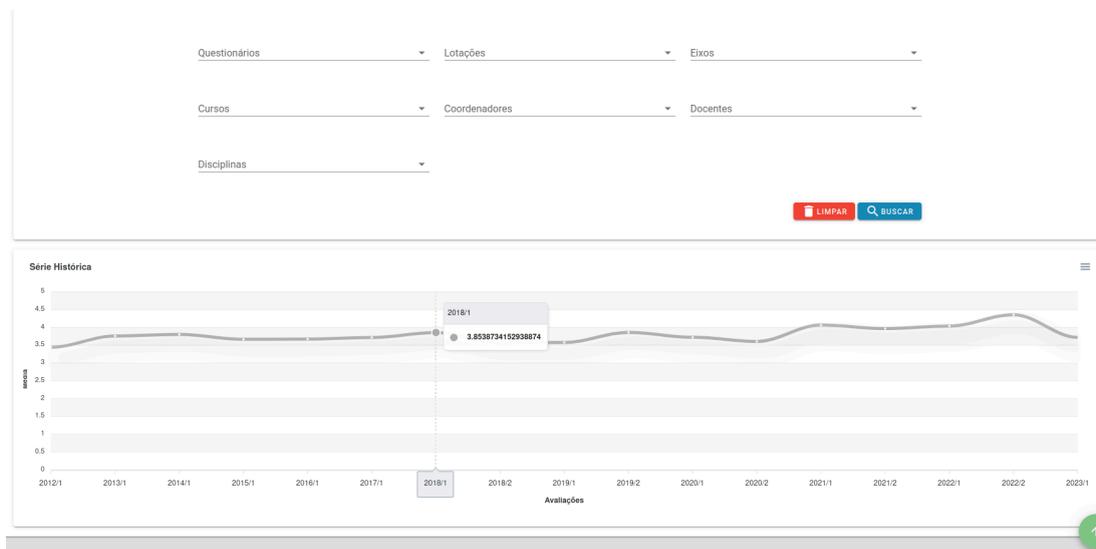
A confecção de um relatório, para que os utilizadores do SIAI possam ter essa informação por meio do sistema, foi observada como uma necessidade pelo autor deste trabalho dada sua importância e sua ausência na atual versão do sistema.

Com intuito de atender esta demanda, foi necessário a criação de um novo serviço no *backend* que, ao ser invocado com opções de filtros em dados disponíveis no sistema, retorna um JSON para o *frontend* contendo as informações solicitadas.

Este JSON, por sua vez, é encaminhado para uma biblioteca de construção de gráficos em **javascript**¹² conhecida como **ApexcharJs**¹³.

O relatório, consiste no cálculo da média obtida sendo distribuída ao longo dos semestres em que a avaliação foi aplicada. Por meio dos filtros de busca, pode-se realizar uma análise específica de acordo com a necessidade do utilizador. A Figura 2 mostra a tela do sistema para emissão do Relatório de Séries Temporais.

Figura 2: Tela do SIAI para visualização do relatório de Análise de série temporal.



Fonte: Autor

4.3. Análise de sentimentos

Para a resolução do problema de análise de sentimentos em questões descritivas do

¹² Linguagem de programação interpretada e baseada em objetos com funções de primeira classe, mais conhecida como a linguagem de script para páginas Web. Fonte: <https://developer.mozilla.org>

¹³ <https://apexcharts.com>

SIAI, a solução adotada foi a utilização de uma versão do modelo BERT, denominado de BERTimbau. As próximas subseções apresentam uma breve conceitualização do BERT, como foi realizado o seu ajuste fino e qual foi o procedimento adotado para integrar seu funcionamento junto ao SIAI. Também será apresentada algumas métricas do desempenho obtido com o modelo.

4.3.1. BERT

O BERT é um modelo de Rede Neural Profunda que pertence à família de modelos *Transformers*, que são arquiteturas de aprendizado de máquina baseadas em mecanismos de atenção. Este modelo introduziu uma abordagem bidirecional para o treinamento, permitindo a absorção do contexto de uma palavra não apenas em relação a suas palavras vizinhas à esquerda, mas também à direita. Essa capacidade bidirecional torna o BERT mais eficiente na compreensão de nuances contextuais e relações semânticas complexas em frases (Devlin et al., 2018).

Este trabalho utilizou o BERTimbau, que é o mesmo modelo do BERT original, porém é pré-treinado com corpus de texto em português.

O BERTimbau, assim como outros modelos de processamento de linguagem natural, faz uso de um pré-processamento conhecido como tokenização, cujo objetivo é transformar o texto em um formato de entrada aceito pela sua rede neural.

Trata-se de um modelo pré-treinado com uma grande quantidade de texto não rotulado extraído de várias fontes, como a web por exemplo. Esta fase de pré-treino do modelo é considerada como “não supervisionada”, já que o texto não possui rótulo.

É importante considerar que, ao descarregá-lo da plataforma em que se encontra hospedado, o modelo já vem com os pesos ajustados por essa fase de pré-treino.

A partir do modelo pré-treinado, é possível a realização de um *fine-tuning*, ou seja, um ajuste para realização de tarefas específicas, como é o caso da análise de sentimentos. Esse ajuste é considerado um aprendizado “supervisionado” e significa que o aprendizado, nesta fase, deve ser realizado com textos rotulados segundo o contexto do problema a ser abordado.

Este texto rotulado deve ser obtido, preferencialmente, do contexto em que o modelo será aplicado e deve ser classificado por um supervisor, um agente que aponte a correta classificação dos dados. A subseção seguinte apresenta o *fine-tuning* do modelo.

4.3.2. Fine-tuning do BERT

Para realizar o *fine-tuning* para o contexto do SIAI foi necessário criar um conjunto de dados, denominado *Dataset*, extraído diretamente da base de dados do sistema, composto por uma amostra de 3 mil registros de respostas descritivas dos respondentes, obtidas aleatoriamente de um total de pouco mais de 208 mil respostas registradas desde o ano de 2012, e por um rótulo, correspondente às classificações de cada uma destas respostas.

Um ponto importante a ser mencionado referente ao *Dataset* é que, como o SIAI garante o anonimato dos respondentes, não é possível relacionar as respostas utilizadas no *fine-tuning* aos seus respectivos respondentes.

Outro detalhe importante referente ao *Dataset* é que este foi previamente limpo antes de sua utilização. Ou seja, as respostas que não puderam ser enquadradas em algum dos rótulos propostos foram descartadas. Foi pensado na possibilidade de adicionar um novo rótulo de neutralidade, porém a quantidade de respostas encontradas com essa característica foi pequena (em torno de 50 respostas). Por isso, a decisão foi o descarte das mesmas.

O *Dataset* obtido foi dividido em três partes, sendo cada uma destas destinadas a uma etapa específica do *fine-tuning*. Do conjunto completo, 80% dos dados foram utilizados para uma etapa de ajuste dos pesos da rede (treino). O restante foi dividido ao meio, onde uma das partes foi utilizada para avaliação de aprendizagem e critério de parada de treino, e a outra parte foi utilizada para testar o desempenho do melhor modelo obtido após o treinamento.

Em posse do *Dataset*, o primeiro passo foi descarregar o modelo pré-treinado da plataforma *Hugging Face*, juntamente com o *tokenizer* recomendado em sua documentação. Para realização do treinamento, deve-se carregar o modelo em em uma Unidade de Processamento Gráfico (GPU) com suporte a CUDA (*Compute Unified Device Architecture*).

Para que o modelo atenda a demanda do problema, deve-se atentar à quantidade de rótulos (denúncia, elogio, reclamação e sugestão) a serem utilizados. Neste caso, dado o contexto do problema, deve-se definir uma camada linear de saída para o modelo de 4 (quatro) neurônios.

Como etapa final de configuração antes de executar o treino da rede neural em si, deve-se definir uma taxa de aprendizagem, também chamada de *learning rate*. A *learning rate* é um hiper-parâmetro que determina a magnitude dos ajustes feitos nos pesos da rede. Ela é utilizada juntamente com um algoritmo de otimização, que tem o objetivo de ajustar os pesos de modo a minimizar uma função de perda (*Loss*).

O *loop* de treino do modelo é composto das seguintes etapas descritas a seguir:

- **Treino:** Nesta seção do *loop*, a rede neural é submetida aos dados de treinamento. A cada iteração, o conjunto de treino é passado para a rede (*forward*), cujo retorno é então passado para a função *Loss* juntamente com os valores verdadeiros. Então, a função *Loss* calcula a perda da predição e atualiza os parâmetros da rede (*backward*). A atualização dos parâmetros pode ser considerada como o aprendizado da rede.
- **Avaliação:** Esta seção utiliza um conjunto de dados do *Dataset* separado para avaliar o modelo, cujo propósito é descobrir se a rede neural está de fato aprendendo. O cálculo da *Loss* é realizado da mesma forma como no treino, com a diferença que os parâmetros da rede não são atualizados.
- **Controle:** Esta seção faz a comparação entre a *Loss* de avaliação atual e a anterior. Caso a atual seja menor que a anterior, significa que a rede aprendeu e o valor anterior é atualizado. Ao perceber que houve um aprendizado, o modelo atual é considerado o melhor modelo até o momento e então é salvo.

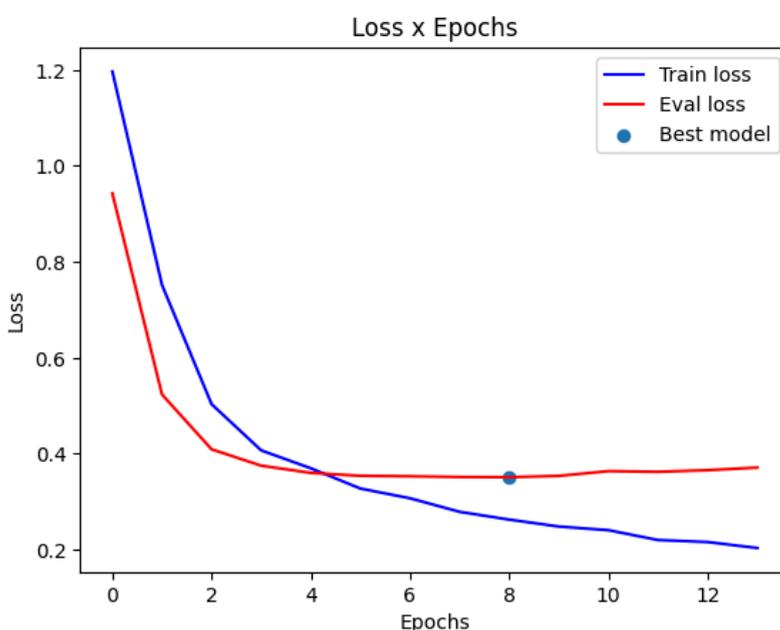
Ao realizar todos estes passos uma vez, dizemos que uma época do treino foi executada. A cada época, é realizada uma medição para saber se a rede neural está

aprendendo. Essa medição é feita por meio de uma simples comparação da *Loss* de avaliação da época atual em relação a *Loss* da época anterior.

Ao executar o *loop* de treino, percebeu-se que o modelo conquistou uma rápida curva de aprendizagem, partindo de uma *Loss* de aproximadamente 1,2 e convergindo para 0,4 em apenas 4 épocas, como mostra a Figura 3.

A partir da quarta época, o modelo iniciou o processo de *overfitting*. Um estado em que o modelo passa a se ajustar sobre o conjunto de treino e não evolui sobre o conjunto de avaliação. Isso significa que, a partir deste ponto, o modelo pode não conseguir generalizar bem para dados novos e não vistos.

Figura 3: Gráfico do *fine-tuning*. *Loss* x Épocas.



Fonte: Autor.

O *loop* de treino ainda percebeu melhorias no modelo após a quarta época. Foram evoluções muito pequenas, porém o suficiente para salvar o melhor modelo na oitava época, encerrando o treino por não evoluir mais sobre o conjunto de avaliação.

4.3.3. Medição do desempenho do modelo

Após o *fine-tuning*, foi realizado um teste para medir o desempenho do modelo que é apresentado nesta subseção. Para isso, foi utilizado um subconjunto do *Dataset* específico para testagem do modelo.

Esse subconjunto refere-se a dados de respostas descritivas do SIAI ao qual o modelo ainda não teve contato. A ocultação desses dados em relação ao *fine-tuning* é realizada para não enviesar o teste do modelo e obter uma métrica mais precisa de como este irá se comportar quando colocado em produção.

Para a realização desta medição foi necessário obter as previsões do modelo sobre o conjunto de testes.

Figura 4: Matriz de confusão

VERDADEIROS	Denúncia	A_D 50	E_{DE} 0	E_{DR} 10	E_{DS} 0
	Elogio	E_{ED} 0	A_E 56	E_{ER} 3	E_{ES} 1
	Reclamação	E_{RD} 11	E_{RE} 5	A_R 98	E_{RS} 6
	Sugestão	E_{SD} 0	E_{SE} 1	E_{SR} 6	A_S 53
	PREDITOS	Denúncia	Elogio	Reclamação	Sugestão

Fonte: Autor

Com os valores verdadeiros e preditos do conjunto de testes, foi possível montar a Matriz de Confusão apresentada na Figura 4 e com isso pode-se estabelecer algumas métricas, segundo Jurafsky e Martin (2019), para calcular a eficácia do modelo.

Para entender a Matriz de Confusão considere que o “A” representa os acertos na predição e o “E” indica que o modelo errou a predição. Juntamente com os indicadores de acertos e erros, em sobrescrito, a matriz informa a letra inicial da classe correspondente.

Por exemplo, o valor $A_E = 56$ indica que o modelo acertou 56 itens da classe “Elogio”, enquanto que o valor $E_{RD} = 11$ indica que o modelo errou 11 itens afirmando que a classe seria “Reclamação”, mas o correto seria “Denúncia”.

Com o auxílio da Matriz de Confusão é possível calcular a eficácia do modelo utilizando as fórmulas apresentadas a seguir, segundo Jurafsky e Martin (2019), onde “R” representa o conjunto de classes e “c” representa a classe em que deseja-se calcular a métrica.

$$(1) \textit{Precision} = \frac{A_c}{A_c + \sum_i E_{ic}}$$

Medida de qualidade que indica a proporção de instâncias classificadas como positivas que são realmente positivas.

$$(2) \textit{Recall} = \frac{A_c}{A_c + \sum_i E_{ci}}$$

Métrica de avaliação que indica a proporção de instâncias positivas que foram corretamente identificadas pelo modelo.

$$(3) \textit{F1 Score} = \frac{2 * \textit{Precision} * \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

Média harmônica da precisão e da recall. A média harmônica dá mais peso aos valores menores, o que significa que a F1 Score será alta apenas se tanto a precisão quanto a recall forem altas.

$$(4) \textit{Accuracy} = \frac{\sum A}{\sum A + \sum E}$$

Métrica de avaliação que mede a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões feitas.

Felizmente, não é necessário realizar todos estes cálculos manualmente. Para obter todos estes resultados foi utilizado a função **classification_report** da biblioteca scikit-learn¹⁴. O resultado da função é apresentado na Figura 5.

Figura 5: Relatório de classificação.

	precision	recall	f1-score	support
DENUNCIA	0.82	0.83	0.83	60
ELOGIO	0.90	0.93	0.92	60
RECLAMACAO	0.84	0.82	0.83	120
SUGESTAO	0.88	0.88	0.88	60
accuracy			0.86	300

Fonte : Autor.

Como pode-se observar na Figura 5, o modelo obteve uma acurácia de 86% de acerto. A coluna support indica a quantidade de exemplos utilizados para cada classe, sendo que o total foi de 300 exemplos.

4.3.4. Integração com o SIAI

O próximo desafio foi como realizar a integração com o SIAI, que é um sistema que utiliza o *framework* Spring Boot em Java, com o modelo de processamento de linguagem natural implementado em Python.

A solução adotada para este problema foi a criação de um projeto de software

¹⁴ <http://scikit-learn.org>

em Python que disponibilizasse uma API Rest a ser consumida pelo *backend* do SIAI.

Para criação desta API Rest foi utilizado o FastApi¹⁵, um *framework* em Python para desenvolvimento de soluções Rest e microsserviços que possui uma rápida curva de aprendizagem e uma excelente performance.

Utilizando o FastApi, foi criado o *endpoint* **/classificar_textos**, que recebe uma lista de respostas descritivas e retorna uma lista contendo a classe juntamente com o **id** da resposta.

Esse *endpoint* é, então, utilizado pelo *backend* do SIAI que, ao invocá-lo, salva as classificações preditas pelo modelo no mesmo banco de dados em que as respostas descritivas estão armazenadas.

Para que o *backend* do SIAI invoque o serviço desenvolvido em Python, foi criado uma nova funcionalidade, disponibilizada na tela do relatório de respostas descritivas do sistema, em que, ao gerar o relatório de respostas, o usuário pode requisitar ao sistema que gere a classificação das respostas obtidas no relatório.

Nas opções de filtros para o relatório, também foi disponibilizado uma opção para filtrar pela classificação da resposta. Com esta opção, o usuário poderá priorizar a conferência das respostas por uma determinada classe. Por exemplo, priorizar a averiguação do que a rede neural classificou como denúncia, antecipando assim, as ações necessárias para corrigir problemas mais importantes.

5. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho embarcamos em uma jornada com o objetivo de melhorar o Sistema de Avaliação Institucional da UFMS, o SIAI. Para que esse objetivo fosse alcançado, várias tarefas foram executadas ao longo do percurso. Esta seção apresenta um resumo de cada uma destas tarefas, bem como uma discussão sobre elas, no qual serão abordados, principalmente, os desafios encontrados e outras possíveis soluções. Também apresenta novas perspectivas para melhorar ainda mais o sistema, sobre a forma de trabalhos futuros.

No início, a RSL foi pensada com o objetivo de confrontar os resultados obtidos na etapa de coleta de requisitos junto a CPA, porém percebeu-se que muitas outras melhorias, além daquelas demandadas pela comissão, poderiam ser testadas e implementadas no sistema.

Até o momento do levantamento de requisitos junto à CPA, o trabalho estaria restrito ao desenvolvimento do Relatório de Ações e o Analisador de Sentimentos. No entanto, a RSL revelou muitas outras possibilidades e, para este trabalho, resolvemos incluir outro requisito, que foi o Relatório de Séries Temporais.

Ao final do levantamento, chegou-se a uma nova lista de requisitos que foram apontados durante a RSL. Como não houve tempo hábil para atender a todos os requisitos levantados, os que não foram desenvolvidos neste trabalho serão indicados como trabalhos futuros.

¹⁵ <https://fastapi.tiangolo.com>

O relatório de ações é considerado como o requisito que iniciou o desenvolvimento deste trabalho, ou seja, foi a principal melhoria solicitada pela CPA. Trata-se de um relatório de análise de dados utilizado para fazer apontamentos de melhorias na instituição.

Com a implementação deste relatório diretamente no SIAI, os usuários terão a possibilidade de obter essas informações de maneira mais eficiente. Além disso, passa-se a ter uma maior confiabilidade nas informações contidas no relatório, pois a forma como era feita antes havia muito mais chances de erro humano ao gerar o relatório pela planilha.

Houve também uma maior flexibilização na emissão do relatório, sendo possível filtrar pela mediana desejada. Algo bem mais complexo para se obter utilizando o método antigo.

Além do relatório de ações, o relatório de séries temporais, um requisito extraído da RSL, também foi implementado neste trabalho de evolução do sistema devido à importância que suas informações fornecem aos usuários.

Trata-se de um relatório para visualizar graficamente as evoluções durante um determinado período em que a avaliação institucional foi aplicada. Seu objetivo é fazer apontamentos se houve melhora ou não entre os anos e semestres em que a avaliação institucional foi realizada.

Em posse das informações sobre as evoluções, ano a ano, é possível saber onde as tentativas de melhorar a instituição estão dando certo e onde estão errando e, com isso, realizar os devidos ajustes para que a melhoria seja contínua.

Mas além deste relatório, a RSL revelou outras análises que podem ser realizadas sobre os dados da avaliação. Análises estas que, ao serem implementadas, podem potencializar ainda mais a iniciativa de tomada de decisão pelos usuários. Dentre essas análises destacam-se: Análise Fatorial; Análise de Variância; Análise de Cluster.

Porém, como dito, não foi possível abordar todas as melhorias a serem aplicadas ao SIAI. Estas análises também ficarão como sugestão para trabalhos futuros.

Apesar destas análises terem ficado como sugestão, a análise de sentimentos foi abordada neste trabalho, sendo criado um analisador de sentimentos para resolver o problema da classificação de respostas textuais do SIAI.

Para este fim, foi utilizado o BERTimbau, um modelo de rede neural derivado de *Transformers* que é utilizado para resolver problemas de processamento de linguagem natural.

O BERTimbau é pré-treinado com corpus de texto em português, sendo necessário apenas a realização de um *fine-tuning* para o contexto do problema da análise de sentimento para tornar sua utilização viável para a tarefa.

O *fine-tuning* foi realizado utilizando um *Dataset* construído a partir das respostas descritivas obtidas por meio dos questionários aplicados pelo SIAI. A utilização de dados pertencentes ao contexto é necessária para evitar que o modelo tenha um comportamento diferente quando posto em produção.

Após o *fine-tuning*, foram feitas medições de desempenho do modelo. As medições apontaram uma acurácia de 86% em dados separados para testar o modelo.

O modelo foi então integrado ao SIAI por meio de disponibilização de uma API Rest criada na mesma linguagem de programação em que o modelo foi desenvolvido. Esta API é consumida pelo SIAI, gerando as classificações para as respostas descritivas.

O problema da análise de sentimentos foi solucionado. Porém, não foram testados outros métodos e modelos, até mesmo alguns apontados na RSL, para verificar se seriam melhores para o contexto do problema. Assim, esta também pode ser uma oportunidade para ser vista em trabalhos futuros.

Um detalhe importante que vale ressaltar refere-se aos dados da classificação, nos quais é possível um revisor marcar se a classificação está correta, e até mesmo podendo ser corrigida caso o revisor julgue necessário. Isso pode auxiliar na criação de um *Dataset* maior para retreinar o modelo, possibilitando aumentar seu desempenho.

Apesar de que nem tudo o que foi levantado como requisitos de melhorias para o SIAI foi desenvolvido, pode-se considerar que o sistema se tornou uma ferramenta melhor, capaz de auxiliar muito mais seus utilizadores na tomada de decisão em relação aos dados contidos em seu banco de dados.

O levantamento de requisitos realizado por este trabalho, apesar de não terem sido todos implementados, pode contribuir para que trabalhos futuros possam ser realizados para melhorar o sistema, além de contribuir para nortear outras instituições a melhorarem seus respectivos sistemas de avaliação sem precisar realizar todo o levantamento novamente.

REFERÊNCIAS

- Arruda, J. A. D., Paschoal, T., & Demo, G. (2019). Uso dos resultados da autoavaliação institucional pelos gestores da Universidade de Brasília. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 24, 680-698.
- Barbisan, A. O. (2002). Avaliação institucional da extensão: conceitos, níveis, parâmetros e indicadores. *Avaliação: Revista da Avaliação da Educação Superior*, 7(3).
- Barbosa, M. E., Maas, H., & Pereira, B. H. (2019). Análise de sentimento por meio de redes neurais artificiais em sistemas de avaliação de cursos de graduação. <https://repositorio.uninter.com/handle/1/1390>
- Erasmio, E. A. L., Duarte, M. M. S. L., de Lucena Pires, E. B. L., & Mendes, R. N. M. (2018). AVALIAÇÃO INSTITUCIONAL: uma análise de indicadores de desempenho institucional em uma IFES. *Revista Observatório*, 4(6), 845-877.
- Farias, A. M. L. D., & Laurencel, L. (2000). Fundamentos de Estatística Aplicada: Módulo I: Estatística Descritiva. *Rio de Janeiro/RJ: Universidade Federal Fluminense*.
- Hoffmann, C., Zanini, R. R., Corrêa, Â. C., Siluk, J. C. M., Schuch Júnior, V. F., & Ávila, L. V. (2014). O desempenho das universidades brasileiras na perspectiva do Índice Geral de Cursos (IGC). *Educação e Pesquisa*, 40, 651-665.
- INEP; INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS

- ANÍSIO TEIXEIRA. Orientações gerais para o roteiro da autoavaliação das Instituições. 2004.
- Jesus, G. R. D., & Bedritichuk, A. G. A. (2018). Autoavaliação institucional: construção e validação de um questionário para o professor. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 23, 628-647.
- JURAFSKY, D.; MARTIN, J. H. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 3rd (Draft). ed. [S.l.]: Prentice Hall, 2019. Disponível em: <<https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>>.
- Konzen, L. Z., & Adams, E. (2016). Autoavaliação institucional: um estudo sobre a análise de dados da autoavaliação em perspectiva diacrônica. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 21, 783-798.
- Luz, R. A. S. D., Petarnella, L., & Silveira, A. (2020). Disponibilidade e Acessibilidade do Sistema de Avaliação Institucional do Centro Estadual de Educação Tecnológica Paula Souza sob a ótica de seus gestores. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 25, 219-237.
- Maba, E. G., & Marinho, S. V. (2012). A autoavaliação institucional no processo de tomada de decisão em IES: Estudo de caso das Faculdades SENAC/SC. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 17, 455-480.
- Monticelli, N. A. M., Rodrigues, S. D., Serafim, M. P., & Atvars, T. D. Z. (2021). Avaliação institucional e gestão estratégica-vínculos necessários para o desenvolvimento institucional. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 26, 315-342.
- Nascimento, J. C. H. B., Bernardes, J. R., de Sousa, W. D., & Lourenço, R. L. (2016). Avaliação institucional: aplicação da Teoria da Resposta ao Item para avaliação discente em Ciências Contábeis. *Future Studies Research Journal: Trends and Strategies*, 8(2), 117-143.
- Nunes, E. B. L. D. L. P., Duarte, M. M. S. L. T., & Pereira, I. C. A. (2017). Planejamento e avaliação institucional: um indicador do instrumento de avaliação do SINAES. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 22, 373-384.
- Paula, M. D. F., & de Azevedo, M. D. (2004). A avaliação institucional na universidade: o caso UFF. *Avaliação: Revista da Avaliação da Educação Superior*, 9(3).
- Pasini, A., Estevez, E., & Pesado, P. (2019, June). Assessment Model for Digital Services provided by Higher Education Institutions. In *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 468-477).
- Perim, G. L., Sakai, M., Almeida, M., Marchese, M., & Matsuo, T. (2008). A avaliação institucional no curso de Medicina da Universidade de Londrina: uma experiência inovadora. *Revista Brasileira de Educação Médica*, 32, 217-229.
- Reis, C. Z. T., Silveira, S. D. F. R., & Ferreira, M. A. M. (2010). Autoavaliação em uma instituição federal de ensino superior: resultados e implicações. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 15, 109-129.
- Shriberg, M. (2002). Institutional assessment tools for sustainability in higher education: strengths, weaknesses, and implications for practice and theory. *Higher education policy*, 15(2), 153-167.
- Silva, J. R. C., Mussi, C. C., Casagrande, J. L., & de Lima, M. A. (2019). A incidência

- dos mecanismos isomórficos nos processos de autoavaliação institucional. *Navus-Revista de Gestão e Tecnologia*, 9(3), 173-186.
- Souza, F., Nogueira, R., & Lotufo, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In *Intelligent Systems: 9th Brazilian Conference, BRACIS 2020, Rio Grande, Brazil, October 20–23, 2020, Proceedings, Part I 9* (pp. 403-417). Springer International Publishing.
- Sousa, R. N. D. (2019). Minerjus: Solução de apoio à classificação processual com uso de Inteligência Artificial. <https://umbu.uft.edu.br/bitstream/11612/1446/1/Rogério%20Nogueira%20de%20Sousa%20-%20Dissertação.pdf>
- Sousa, T. E., & Lambert-de-Andrade, N. (2022). A autoavaliação institucional e sua contribuição para as tomadas de decisões democráticas. *Vértices (Campos dos Goitacazes)*, 24(1).
- Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune bert for text classification?. In *Chinese Computational Linguistics: 18th China National Conference, CCL 2019, Kunming, China, October 18–20, 2019, Proceedings 18* (pp. 194-206). Springer International Publishing.
- Tavares, D. A., Rosa, M. J., & Amaral, A. (2010). Does the EUA institutional evaluation programme contribute to quality improvement?. *Quality Assurance in Education*.