Predição de Desistência em Turmas de Programação Utilizando Sentimentos

Rodrigo N. F. Baessa, UFSC-Joinville, rnfbaessa@gmail.com, ORCID:0009-0005-5564-2291
Carolina B. Wang, UFSC-Joinville, carolina.wang.ufsc@gmail.com, ORCID:0009-0006-1387-7348
Ricardo J. Pfitscher, UFSC-Joinville, ricardo.pfitscher@ufsc.br, ORCID:0000-0003-0539-2237
Benjamin G. Moreira, benjamin.grando@ufsc.br, UFSC-Joinville, ORCID:0000-0002-0339-4012
Tatiana R. Garcia, UFSC-Joinville, tatiana.garcia@ufsc.br, ORCID:0009-0001-3571-915X
Luiz C. Camargo, Católica de SC, lzcamargo@outlook.com, ORCID:0000-0001-7879-9893

Resumo: A desistência precoce em disciplinas de programação é um problema frequentemente enfrentado em cursos de graduação. Este artigo avalia o quanto o sentimento dos estudantes da turma pode auxiliar na predição de desistência. Foram utilizados dados de avaliações de desempenho e sentimentos dos estudantes com o uso de métodos de aprendizado de máquina. A análise evidenciou diferenças no desempenho de alunos concluintes, sugerindo a eficácia das variáveis propostas na construção de modelos preditivos, principalmente com a adição do atributo referente aos sentimentos. O estudo contribui para a identificação de estudantes desistentes, permitindo a implementação de intervenções precoces e oferecendo uma abordagem para a predição dessas desistências. Os resultados mostram que a adição da informação de sentimento aumentou em 2% a sensibilidade da classificação e até 30% a precisão dos modelos preditivos.

Palavras-chave: Desistência estudantil, aprendizagem de máquina, análise de sentimentos.

Predicting Dropout in Programming Classes Through Sentiment Analysis

Abstract: Early dropout in programming courses is a common problem in undergraduate courses. The present research evaluates how students' sentiments can help predict dropout. Data from performance assessments and student sentiments were analyzed using machine learning methods. The analysis showed differences in the performance of graduating students, suggesting the effectiveness of the proposed variables in building predictive models, especially with the addition of the attribute related to sentiment. The study contributes to identifying dropout students, allowing the implementation of early interventions, and offering an approach to predicting dropout in educational contexts. The results show that adding sentiment information increased the classification sensitivity by 2% and the precision of predictive models by up to 30%.

Keywords: Student dropout, machine learning, sentiment analysis.

1. Introdução

A evasão de estudantes no ensino superior é tema de crescente preocupação e estudo. Esse fenômeno não apenas impacta indivíduos, mas também exerce influência significativa sobre o progresso educacional, econômico e social de uma sociedade (Nagai e Cardoso, 2017). Ainda, segundo Nagai e Cardoso (2017), a evasão não é apenas um problema localizado, mas sim um desafio global com padrões comportamentais relativamente consistentes em certas áreas de estudo, apesar das variações entre instituições e contextos socioeconômicos e culturais.

Para os estudantes, a evasão pode acarretar consequências profundas. Além das dificuldades financeiras associadas à interrupção dos estudos, como custos perdidos com mensalidades e investimentos anteriores, há também impactos psicológicos significativos (Mallada, 2011). Sentimentos de fracasso e inadequação podem emergir, afetando a autoestima e o bem-estar emocional dos estudantes, potencialmente levando à

desmotivação prolongada ou até ao abandono completo da educação. Do ponto de vista da sociedade, a evasão representa uma perda de capital humano valioso (Gaioso, 2005). Assim, altas taxas de evasão podem reduzir a competitividade nacional e o progresso coletivo, impedindo o crescimento sustentável e a prosperidade geral.

Durante a realização de um curso de graduação, percebe-se que diversos alunos não conseguem completar o cronograma de determinadas disciplinas por vários motivos, como falta de motivação e dificuldade em compreender o conteúdo. A não conclusão da matéria configura um tipo de fracasso escolar: a desistência (Nagai e Cardoso, 2017). No campo acadêmico, há uma crescente adoção de métodos analíticos e baseados em Inteligência Artificial (IA) para prever e mitigar a evasão e a desistência de alunos (Macarini *et al.*, 2019; Queiroga *et al.*, 2020; Morais *et al.*, 2018; Souza e Moreira, 2024b; Costa *et al.*, 2020; Kim; Yoo e Kim, 2023; Matz *et al.*, 2023). Normalmente, esses esforços de pesquisa treinam algoritmos de classificação com métricas coletadas passivamente, como desempenho acadêmico nas atividades e frequência. No trabalho aqui apresentado, investiga-se como o sentimento dos estudantes em uma turma pode ser utilizado para auxiliar na previsão da desistência estudantil.

O presente estudo contribui com a elaboração de um modelo de classificação integrando abordagens passivas e ativas, a fim de detectar previamente indicadores de potencial evasão, consequentemente, colaborar com o desenvolvimento de pesquisa no campo acadêmico, especificamente em relação à evasão estudantil. Para avaliação da proposta, foram utilizados dados de turmas de programação de computadores, que são disciplinas dificuldades que envolvem a quebra do paradigma do raciocínio puramente matemático para o pensamento lógico iterativo de uma máquina computacional, com altas taxas de desistência e reprovação (Pereira; Seabra e Souza, 2021; Moreira *et al.*, 2024). Em suma, este trabalho mostra que a adição da informação de sentimento da turma melhorara em torno de 2% a sensibilidade da classificação e até 30% a precisão dos modelos baseados em Inteligência Artificial (IA) elaborados neste trabalho.

O restante deste artigo está organizado da seguinte forma. Na Seção 2, trabalhos relacionados são apresentados. Na Seção 3, são apresentas a metodologia de análise, com detalhes sobre as turmas analisadas, as técnicas de predição explícitas e implícitas. Por fim, na Seção 4, discute-se os resultados e na Seção 5, estão as conclusões deste trabalho.

2. Trabalhos Relacionados

Evasão estudantil é reportada em vários trabalhos, sendo na Tabela 1 trazido cinco trabalhos com aspectos próximos a este. Esses trabalhos endereçam permanência ou predição de desistência, com sua maioria incluindo Análise de Sentimentos como um método de análise de dados, mas não como atributo na avaliação dos modelos baseados em Aprendizagem de Máquina. A adoção desta estratégia é considerada promissora neste trabalho, com bom desempenho na avaliação de atributos sob o método *class_feeling*.

Quanto ao aspecto Técnicas de Inteligência Artificial (IA), diferentes técnicas são utilizadas na avaliação dos modelos por esses trabalhos (Tabela 1), sendo os trabalhos de Kim, Yoo e Kim (2023) e Ananthi e Arul (2023) com quatro técnicas, e em comum SVC (*Support Vector Classification*) e RF (*Randon Forest*). Ao passo que neste trabalho, seis técnicas de IA, tais como: DT (*Decision Tree*), NB (*Naive Bayes*), RF (*Randon Forest*), MLP (*Multilayer Perceptron*) LR (*Logistic Regression*) e SVM (*Support Vector Machines*) utilizadas na elaboração do modelo de classificação, além de PLN, amplificando campo experimental.

Coleta de dados é um dos aspectos essenciais para elaboração de modelos, sendo que no trabalho de Pant (2023) são utilizados os métodos ativo e passivo como meio

Tabela 1. Aspectos dos Trabalhos Relacionados

Trabalho	Fonte de Dados	Métodos de Coleta de Dados	Técnicas de IA	Análise de Sentimentos	Endereçamento
(Pfitscher *et al.*, 2023)	Comentários de estudantes	Ativa	PLN/LSTM	Sim	Permanência estudantil
(Pant, 2023)	Surveys Entrevistas	Ativa e Passiva	Ferramenta NVivo	Sim	Retenção MOOCs
(Bóbó *et al.*, 2022)	Ambiente Virtual	Passiva	Lexical FrameNet	Sim	Evasão Estudantil
(Ananthi e Arul, 2023)	Ambiente Virtual	Passiva	RF, SVC MNB e LR	Sim	Evasão Estudantil
(Kim; Yoo e Kim, 2023)	Dados Acadêmicos	Passiva	SVC,DT RF e KNN	Não	Predição de Desistência

de coleta de dados, assim como neste trabalho. Ao contrário do método passivo, o método ativo requer o engajamento de indivíduos em contribuir com o método, neste caso os estudantes, limitando o acesso à fonte de dados e a quantidade. Assim, percebe-se que apenas dois (Pfitscher *et al.* (2023) e Pant (2023)) dos cinco trabalhos relacionados utilizam o método ativo de coleta de dados.

3. Metodologia

A metodologia utilizada é baseada na integração de dados provenientes de duas fontes distintas, e posteriormente uma análise do conjunto de dados resultante visando avaliar indicadores que possam sinalizar a possibilidade de desistência de estudantes matriculados na disciplina. Os dados coletados para a realização das análises deste trabalho são obtidos por meio de dois processos distintos: ativo e passivo.

O primeiro, método ativo, refere-se a coleta de sentimentos dos estudantes com relação à disciplina via questionários (detalhes deste processo estão disponíveis em Pfitscher *et al.* (2023)). No segundo, método passivo, são coletadas informações referentes ao desempenho do estudante ao longo do semestre, incluindo presenças e notas no Moodle (detalhes sobre como são extraídos esses dados no ambiente Moodle podem ser encontrados em Souza e Moreira (2024a)). Para mensurar o impacto dos sentimentos em relação à disciplina com a predição da desistência dos alunos nas turmas, inicialmente foram treinados modelos que utilizaram apenas os dados coletados de forma passiva (presença, notas e atividades finalizadas). Posteriormente, foi adicionada a informação sobre sentimento da turma para utilização dos modelos. O código-fonte e os dados anonimizados utilizados nos modelos estão disponíveis no GitHub.*

3.1. Turmas analisadas

São utilizados dados provenientes de turmas de disciplinas de programação de dois semestres distintos. Em ambas as turmas o professor responsável fez a marcação manual dos alunos desistentes e não desistentes de suas turmas. As disciplinas são ministradas em dois encontros presenciais semanais. A Turma 1 possui 83 alunos matriculados e um total de 26 atividades avaliativas aplicadas.† Conforme o relatório de presença foram anotadas presenças em 26 seções/aulas e foram obtidas 28 respostas nos questionários sobre sentimentos. A Turma 2 possui 93 alunos e foram realizadas 25 atividades avaliativas. Foi anotada a frequência em 29 encontros e foram registradas 43 respostas aos questionários de sentimento.

*Repositório com os dados disponível em https://github.com/MecatronicaUFSC/TCCs/tree/main/RodrigoBaessa

†As turmas foram elaboradas visando coletar dados sobre a realização de atividades. As atividades são questionários e Laboratórios Virtuais de Programação aplicados periodicamente.

3.2. Dados utilizados

A aquisição dos dados referentes aos sentimentos dos alunos com relação à disciplina foi realizada de forma *ativa*, por meio de questionários com perguntas abertas respondidos de forma anônima, assim como realizado no estudo de Pfitscher *et al.* (2023).

Uma das perguntas é a seguinte: "Qual é o seu sentimento em relação à disciplina?". Dessa forma, as respostas são textos elaborados livremente pelos estudantes que descrevem seu sentimento perante a disciplina em questão. No entanto, algumas frases podem possuir termos relacionados a sentimentos e outras não, bem como, conter emoções que podem causar confusão e incertezas no momento de identificar o real sentimento do estudante. Após a coleta das respostas, foi realizada manualmente uma classificação psicopedagógica, que empiricamente busca distinguir termos atribuídos aos sentimentos dos demais em cada frase obtida.

O primeiro passo desta classificação é a definição de categorias, listam-se posteriormente os itens sobre aspectos positivos e negativos na experiência com o curso/disciplina utilizando a análise de conteúdo, e em seguida é realizada uma pré-análise flutuante construindo categorias de classificação. Maiores detalhes sobre a distribuição das respostas podem ser encontrados em Pfitscher *et al.* (2023). Os dados de desempenho do estudante (coleta passiva) são coletados por meio de dois relatórios do ambiente Moodle, sendo estes relatórios de presença e notas. Estas escolhas foram feitas de modo a não restringir o modelo às instituições que utilizem o Moodle, uma vez que essas informações podem estar disponíveis também em outras plataformas.

3.3. Pré-processamento de dados

Para assegurar a qualidade e a confiabilidade dos dados obtidos, foi realizada uma etapa de pré-processamento. Esta fase envolve a preparação dos conjuntos de dados que serão utilizados nos testes, assegurando que estejam adequadamente formatados e representem fielmente as condições desejadas para a avaliação. A técnica envolve quatro passos: limpeza, integração, transformação e redução de dados.

Um primeiro passo da etapa de limpeza dos dados é anonimização, sendo atributos utilizados de identificação dos estudante alterados para valores genéricos. Subsequentemente, foram removidos dados irrelevantes para o estudo como, por exemplo, endereço de e-mail, sobrenome, número de matrícula, entre outros atributos similares. No relatório de presença foi selecionada a quantidade de cada tipo de presença atribuída (presente, atrasado e ausente), a quantidade de sessões anotadas, a relação de pontos e a porcentagem de presença. Já no relatório de notas são selecionados os valores das notas atribuídas para cada uma das atividades realizadas pelo estudante.

Subsequente a limpeza de dados, é realizada a integração dos dados dos relatórios extraídos do ambiente Moodle e as respostas obtidas pelos questionários. As respostas com relação aos sentimentos dos alunos obtidos por meio dos questionários abertos não totalizam o número de participantes da turma e nem podem ser atrelados diretamente a um indivíduo, uma vez que os questionários são respondidos de forma anônima.

Na etapa de transformação de dados, criou-se variáveis e classificações que possam ser interpretadas pelos algoritmos de IA baseados nas informações disponíveis nos relatórios e questionários. Converteu-se os dados transformados em atributos individuais do aluno ou coletivo da turma com valores numéricos. Com isso, o conjunto de dados, que antes era dividido em quatro relatórios e estruturas de dados diferentes, é reduzido à 23 variáveis numéricas para cada um dos estudantes. Posteriormente, para cada uma das abordagens, é mostrado seu impacto na classificação da desistência.

3.4. Sistema de pontuação de notas, presença e sentimento da turma

Cada um dos relatórios de desempenho e o relatório de sentimento precisam ser transformados em uma nota que pode ser manipulada e comparada. Para isso, criaram-se sistemas de pontuação baseados nas informações presentes nestes relatórios. Para o cálculo da pontuação referente à presença do discente é considerado o peso para cada uma das aulas anotadas, de maneira que em dias com mais alunos presentes, represente um peso maior para a pontuação final da presença. Este método tem o propósito de mitigar efeitos externos que possam dificultar a presença, sendo a forma de cálculo mostrada na Equação 1.

$$W_i = \frac{Me_i}{Max_i} \tag{1}$$

Conforme mostrado na Equação 1, o peso W atribuído a cada dia i é obtido pela divisão da média das pontuações do Moodle da classe para aquele dia pela pontuação máxima definida pelo responsável da matéria dentro do ambiente do Moodle. Sendo assim, pode-se calcular a presença computada para a turma para cada um dos dias anotados de acordo com a Equação 2, onde P_f corresponde ao valor da pontuação final, Po_i é a pontuação atribuída pelo Moodle e d é o número total de dias anotados no relatório de presença.

$$P_f = \sum_{i=1}^{d} \frac{Po_i \times W_i}{Max_i} \tag{2}$$

O cálculo da pontuação das notas leva em consideração a média da turma na avaliação em questão. Caso o estudante obteve uma nota relativamente mais baixa com relação a média da turma, isso reduz a pontuação final mais do que em avaliações, em que a média da turma foi mais baixa. Esse recurso é utilizado para ponderar e acompanhar o desempenho do aluno conforme o resultado geral da turma, como mostrado na Equação 3, onde $P_f^{'}$ é atribuído a pontuação final das notas, G_{max} é a maior soma de notas possível, G_k representa à nota individual para a atividade k, a maior nota obtida entre a classe na atividade é representada por A_k, a média das notas da classe na atividade é simbolizada por M_k e n é a quantidade total de atividades avaliativas realizadas ao decorrer do semestre.

$$P_f^{'} = \frac{G_{max} - \sum_{k=1}^{n}(1 - \frac{G_k}{A_k}) \times M_k}{G_{max}} \times 100 \tag{3}$$

O aspecto relacionado aos sentimentos, principal enfoque deste trabalho, exige a conversão de dados qualitativos para uma forma de escala similar ao adotado nas demais pontuações. São apresentadas três abordagens distintas de como transformar estas classificações dos sentimentos (positivo ou negativo) em um atributo que possa ser utilizado na previsão de desistentes.

3.4.1. Sistema de pontuação dos sentimentos

O primeiro método para pontuação dos sentimentos é a distribuição de um valor binário de acordo com o sentimento geral das respostas obtidas na turma. São contabilizadas as respostas obtidas e suas classificações, o sentimento que predomina, ou seja, o sentimento mais recorrente dentre todos na turma é utilizado no atributo de sentimento, como mostrado na Equação 4. Conforme a equação, a pontuação atribuída será de 0 em caso do sentimento negativo predominar nas respostas e o valor de 1 em caso

de sentimento predominantemente otimista dos estudantes da turma. Esse método é denominado Método Binário (MB).

$$P''_{f_1} = \beta \times 100 \tag{4}$$

Outro método é a atribuição de um valor numérico que representa a porcentagem de alunos otimistas com a disciplina dentro da turma. Similarmente ao MB, é contabilizada a quantidade de respostas com registros de sentimentos positivos dos estudantes e então é convertida em uma porcentagem com relação ao número de respostas obtidas (Equação 5) e denominada Porcentagem de Otimismo (PO).

$$P''_{f_2} = \frac{n_{positivos}}{n_{respostas}} \times 100 \tag{5}$$

O último método empregado é um passo subsequente ao método PO. Nesse método a porcentagem obtida pelo método anterior tem seu peso modificado de acordo com as demais pontuações adquiridas pelo estudante (presença e notas). Esse recurso visa diferenciar os valores atribuídos ao sentimento dentre os participantes da mesma turma. Com isso, quanto maior for a relação entre a soma das pontuações anteriores e a soma máxima obtida na turma (o que indica que o aluno está indo melhor no semestre), maior será o peso atribuído ao otimismo do estudante. Seguindo esta mesma lógica, quanto menor a relação entre a pontuação somada do estudante e a pontuação máxima da turma, menor será o sentimento de otimismo atribuído ao estudante em questão (Equação 6) e denominada Porcentagem de Otimismo Ponderada (POP).

$$P''_{f_3 i} = \frac{(P_{fi} + P'_{fi})}{S_{max}} \times P''_{f_2} \tag{6}$$

Na Equação 6 pode-se observar o cálculo empregado para a criação da pontuação de sentimento. Assim, o resultado da pontuação para o aluno i é dado por $P''_{f_3 i}$, que calcula-se pela soma das pontuações de presença (P_{fi}) e notas (P'_{fi}), dividido pelo somatório máximo atingido pela turma $Smax$ e multiplicado pela porcentagem de otimismo da turma com relação a matéria P''_{f_2}.

3.5. Modelo de previsão de desistência

Na elaboração do modelo foram usados os atributos discutidos neste artigo, sendo que o atributo que descreve o sentimento da turma é atribuído por meio de um dos três métodos de pontuação descritos na seção 3.4.1: Método Binário (MB), Porcentagem de Otimista (PO) e Porcentagem de Otimismo Ponderada (POP).

Na elaboração do modelo de classificação, foram empregados diferentes técnicas de aprendizado de máquina: Árvore de Decisão, Naive Bayes, Random Forest, Multilayer Perceptron, Regressão Logística e Máquina de Vetor de Suporte. Os algoritmos foram utilizados em sua forma de configuração padrão fornecidas pelas bibliotecas. Com o intuito de testar as diferentes técnicas de aprendizado de máquina e seus resultados, utilizou-se três métodos de divisão de dados em treinamento (parcela dos dados que os modelos utilizarão para treinar) e teste (parcela dos dados em que o modelo será aplicado e avaliado). O método de amostragem aleatória de dados (sendo 67% dos dados para treinamento e 33%para testes); o método *holdout* (67% dos dados para treinar e 33% para testar); e o método de validação cruzada 10-*fold* ($k = 10$).

Para a análise, optou-se por unir as duas turmas em um grupo de dados para o modelo ter uma variedade maior de dados para o treinamento. Devido à seleção

aleatória dos estudantes para os conjuntos de dados, a técnica de amostragem aleatória foi implementada dez vezes em cada análise, tal procedimento visa melhorar a capacidade de representação da base de dados, buscando minimizar as variações naturais à seleção aleatória e promovendo uma amostra mais fidedigna da turma em questão. No método de validação cruzada 10-*fold*, foram definidas 10 divisões.

Tanto na amostragem aleatória quanto na validação cruzada, não é necessário dividir os dados manualmente, basta fornecê-los aos modelos. No entanto, ao utilizar o método *holdout* para dividir os dados, foi necessário assegurar a representatividade dos comportamentos de desistência em ambos os conjuntos. Assim, além de dividir os dados manualmente na proporção de 67% para treinamento e 33% para teste, garantiu-se que a proporção de alunos classificados como desistentes fosse equivalente nos dois grupos.

4. Resultados

Para avaliar os modelos quanto a previsão da desistência escola foi primeiramente avaliado o quanto o atributo de sentimento da turma impacta nos modelos de classificação considerando diferentes métricas de ganho de atributo. Depois, os modelos de classificação foram avaliados conforme as métricas *sensibilidade*, *precisão* e *F1-score*.

Para qualificar a relevância da métrica de sentimento nos modelos de classificação, foi realizada uma análise usando como entrada 23 atributos distintos obtidos da coleta ativa e passiva. Utilizando os métodos de avaliação de atributos, foi possível identificar as características mais importantes para a análise realizada. Na Figura 1 pode-se observar o melhor resultado encontrado do ranking dos 10 melhores atributos obtidos por meio do cálculo do atributo *class_feeling* utilizando a ponderação baseada nas demais pontuações (POP). O novo atributo criado para descrever os sentimentos dos estudantes ficou em 3º lugar no método de seleção de relação de ganho, e em 1° lugar no ganho de informação. Já nos métodos Chi2 e ANOVA, o sentimento da turma não foi considerado entre os 10 mais importantes fatores para a classificação. Estes resultados indicam que o parâmetro referente ao sentimento da turma é relevante para alguns modelos de classificação e para outros não implicará em diferenciação entre as classes. Conforme será discutido em seguida, a classificação realizada pelo modelo Naive Bayes é um dos casos que o sentimento da turma é benéfico para prever a possível desistência.

Rank	Ganho de Informação	Relação de Ganho	Chi2	ANOVA
1	class_feeling	sequencial_missing_22	grades_below_5	grades_below_5
2	grades_between_7_5_10	sequencial_missing_11	grades_average	grades_between_7_5_10
3	grades_below_5	class_feeling	grades_between_0_2_5	grades_average
4	grades_average	activities_complete_majority	grades_between_7_5_10	grades_between_0_2_5
5	grades_between_0_2_5	activities_incomplete_majority	grades_below_mean	grades_below_mean
6	activities_complete_minority	grades_between_7_5_10	important_grades_below_mean	important_grades_below_mean
7	activities_incomplete	grades_below_5	important_activities_complete_majority	important_activities_complete_majority
8	grades_below_mean	grades_average	important_activities_complete_minority	important_activities_complete_minority
9	activities_incomplete_minority	grades_between_0_2_5	important_activities_incomplete	important_activities_incomplete
10	activities_complete_majority	activities_complete_minority	important_activities_incomplete_majority	important_activities_incomplete_majority

Figura 1. Seleção de atributos para a turma segundo ponderação baseada nas demais pontuações

Na avaliação dos modelos de previsão de desistência investigou-se quatro diferentes proporções do período letivo em duas turmas distintas. Embora o uso de 100% dos dados não seja factível para intervenções precoces, ele serve como referência para a melhor métrica. Para simular situações reais, adotaram-se porcentagem de 75%, 50% e 25% dos dados, que objetiva medir o desempenho da classificação para testar a capacidade

do sistema de identificar precocemente uma desistência.

Por questões de espaço, serão apresentados somente os resultados do melhor modelo de classificação entre os testados, o Naive Bayes. A Tabela 2 mostra os valores das métricas de sensibilidade, precisão e F1-Score para cada modelo (AS - Ausência de sentimento, MB - Método Binário, PO - Porcentagem de Otimismo e POP - Porcentagem de Otimismo Ponderada) e proporção do período letivo concluído no momento do teste. As células em destaque na tabela indicam o melhor caso para cada métrica em cada proporção do período letivo concluído.

Tabela 2. Comparação entre os diferentes métodos

Modelo	Prop. do período letivo	Sensibilidade	Precisão	F1-score
AS	100%	98,36	43,90	60,70
	75%	98,1	44,22	60,96
	50%	96,69	**45,92**	62,27
	25%	93,96	43,58	59,54
MB	100%	**100**	33,73	50,44
	75%	**100**	48,11	64,97
	50%	**98,8**	43,96	60,85
	25%	**94,6**	44,49	60,52
PO	100%	**100**	34,28	51,06
	75%	98,57	**48,51**	**65,02**
	50%	98,15	45,19	61,89
	25%	94,12	**48,31**	**63,85**
POP	100%	**100**	**71,79**	**83,58**
	75%	99,64	46,81	63,70
	50%	98,21	45,91	**62,57**
	25%	93,98	45,73	61,52

A análise dos resultados mostra que a inclusão do sentimento implicou em melhoria nas três métricas avaliadas na grande maioria das proporções de conclusão do semestre letivo. A única exceção é a métrica precisão, com 50% de conclusão do semestre, que teve seu melhor resultado com o modelo sem uso de sentimentos, com uma taxa de 45,92%. Ainda sobre a precisão, destaca-se que quase todos os casos apresentaram resultados ruins (abaixo de 50%), com exceção do modelo POP, que atingiu 71,79% utilizando 100% dos dados do semestre. A baixa precisão aliada com os altos valores de sensibilidade, indicam que os modelos tiveram uma alta taxa de falsos positivos, ou seja, os modelos classificaram como desistências casos de não desistência. Por outro lado, é importante destacar os resultados obtidos com a métrica de sensibilidade, que é balizada pelos casos de desistência de fato (proporção de verdadeiros positivos em relação a todos os casos positivos).

O modelo baseado em método binário de sentimento obteve os melhores resultados de sensibilidade para qualquer proporção da turma, oferecendo ganhos na ordem de 2% quando comparado ao modelo com ausência de sentimento. Por mais que em termos relativos 2% seja um valor baixo, em termos absolutos, representa um ou dois casos de desistência que poderiam ter sido mitigados caso houvesse a detecção precoce. Ainda no contexto da desistência precoce, a observação da métrica F1-score mostra que os modelos de Porcentagem de Otimismo (PO) e ponderação das demais pontuações (POP) se mostraram mais equilibrados entre a taxa de acerto dos casos de desistência

(sensibilidade) e a diminuição de falsos positivos (precisão). O modelo PO apresentou o melhor resultado de F1-score com 75% e 25% da conclusão do curso e o modelo POP com 50%. Assim, considera-se que o atributo *class_feeling* é importante na predição de estudantes desistentes, possibilitando a mitigação dos casos precoces identificados.

4.1. Discussões e limitações

Apesar dos resultados promissores observados neste artigo, existem algumas limitações a serem consideradas. Primeiro, a ausência de bases de dados públicas relacionadas ao sentimento de estudantes em disciplinas, obriga desenvolvimento de uma base de dados própria. Por consequência, a quantidade de dados utilizada para treinamento dos modelos de IA é pequena quando comparado com o estado da arte na pesquisa relacionada a inteligência artificial. Contudo, como este é um projeto de pesquisa em andamento, espera-se que novos dados sejam utilizados para validação e ajuste dos modelos em trabalhos futuros.

Outro aspecto que pode limitar os achados científicos desta pesquisa é o viés na análise subjetiva dos sentimentos. Considerando que as respostas dos estudantes foram avaliadas manualmente, é possível que a anotação entre positivo e negativo para determinada frase esteja sujeita a percepção de quem lê. Para minimizar este impacto, esta pesquisa utilizou da avaliação de duas psicopedagogas que avaliaram as respostas de forma independente e isolada. Em caso de não convergência das avaliações, as respostas foram anotadas como contendo os dois sentimentos (positivo e negativo). Para automatizar a classificação se faz necessário empregar ferramentas de Processamento de Linguagem Natural (PLN) para análise dos dados, uma vez que a coleta de sentimentos foi realizada por meio de questionários abertos com respostas livres. Esta análise será objeto de pesquisa em trabalhos futuros.

Por fim, destaca-se que os modelos utilizados neste trabalho foram treinados em sua forma padrão, sem um necessário ajuste de hiperparâmetros. O aprimoramento destes modelos será realizado em trabalhos futuros de acordo com o surgimento de novos dados oriundos da continuidade da pesquisa.

5. Conclusão

A desistência estudantil em turmas de programação tem sido objeto de investigação tanto na academia quanto no mercado do ensino superior. Principalmente quanto na possibilidade de uso de modelos automatizados que possam prever de forma precoce potenciais desistências. Este artigo investigou o uso do sentimento dos estudantes como métrica de entrada em modelos de aprendizagem de máquina para previsão de desistência. Os resultados deste artigo demonstraram a utilidade da métrica em um modelo Naive Bayes na previsão da desistência escolar mesmo com uma quantidade limitada de dados. Com apenas 25% dos dados disponíveis, o modelo alcançou alta sensibilidade (94,12%). Isso destaca a capacidade dos modelos de fornecer previsões precisas desde estágios iniciais do processo educacional. Em análises com 50% dos dados, o modelo continuou a exibir alta sensibilidade (98,21%), esses resultados sublinham a robustez dos métodos de aprendizado de máquina ao longo do período letivo, permitindo adaptações dinâmicas a medida que mais dados estão disponíveis.

Quanto aos trabalhos futuros, a continuidade da presente pesquisa pretende investigar como os sentimentos podem ser classificados de forma automatizada para serem incluídos em modelos dinâmicos de avaliação da turma. Ou seja, pretende-se avaliar se o conjunto de variáveis utilizada para prever a desistência deve ser alterada ao longo do curso e como o sentimento influencia essa classificação.

Referências

Ananthi, C. M. T.; Arul, L. R. P. J. Multifaceted sentiment detection system (msds) to avoid dropout in virtual learning environment using multi-class classifiers. **International Journal of Advanced Computer Science and Applications**, The Science and Information Organization, v. 14, n. 4, 2023.

Bóbó, M. L. *et al.* Using sentiment analysis to identify student emotional state to avoid dropout in e-learning. **International Journal of Distance Education Technologies (IJDET)**, IGI Global, v. 20, n. 1, p. 1–24, 2022.

Costa, A. G.; Queiroga, E.; Primo, T. T.; Mattos, J. C. B.; Cechinel, C. Prediction analysis of student dropout in a computer science course using educational data mining. In: IEEE. **XV Conferencia Latinoamericana de Tecnologias de Aprendizaje (LACLO)**. [S.l.], 2020.

Gaioso, N. P. d. L. O fenômeno da evasão escolar na educação superior no brasil. **Brasília, DF: Universidade Católica de Brasília**, p. 20, 2005.

Kim, S.; Yoo, E.; Kim, S. **Why Do Students Drop Out? University Dropout Prediction and Associated Factor Analysis Using Machine Learning Techniques**. 2023.

Macarini, L. A. B.; Cechinel, C.; Machado, M. F. B.; Ramos, V. F. C.; Munoz, R. Predicting students success in blended learning—evaluating different interactions inside learning management systems. **Applied Sciences**, MDPI, v. 9, n. 24, p. 5523, 2019.

Mallada, F. J. R. La gestión del absentismo escolar. **Anuario Jurídico y Económico Escurialense**, n. 44, p. 579–596, 2011.

Matz, S. C. *et al.* Using machine learning to predict student retention from socio-demographic characteristics and app-based engagement metrics. **Scientific Reports volume 13**, 2023.

Morais, A. M. d. *et al.* Abordagem avaliativa multidimensional para previsão da evasão do discente em cursos on-line. Universidade Federal de Campina Grande, 2018.

Moreira, B. G.; Pfitscher, R. J.; Camargo, L. C.; Garcia, T. R. Análise de sentimentos e emoções com o uso da ferramenta orange datamining: uma avaliação a partir da tradução dos textos do português. **Anais do Computer on the Beach**, v. 15, p. 352–354, 2024.

Nagai, N. P.; Cardoso, A. L. J. A evasão universitária: Uma análise além dos números. **Revista Estudo & Debate**, v. 24, n. 1, 2017.

Pant, e. a. Thematic and sentiment analysis of learners' feedback in moocs. **Journal of Learning for Development Vol. 10 No. 1: ISSUE 1, p. 38-54**, 2023.

Pereira, D. E. F.; Seabra, R. D.; Souza, A. Diniz de. Ferramentas de apoio ao ensino introdutório de programação: um mapeamento sistemático. **Revista Novas Tecnologias na Educação**, v. 18, n. 2, p. 491–500, jan. 2021. Disponível em: ⟨https://seer.ufrgs.br/index.php/renote/article/view/110289⟩.

Pfitscher, R. J. *et al.* Análise de sentimentos em turmas de programação com vistas ao apoio à permanência estudantil. **Congresso Brasileiro de Informática na Eduacação**, 2023.

Queiroga, E. M. *et al.* A learning analytics approach to identify students at risk of dropout: A case study with a technical distance education course. **Applied Sciences**, MDPI, v. 10, n. 11, p. 3998, 2020.

Souza, B. D. d.; Moreira, B. G. Detecção de desistência de estudantes em disciplinas ofertadas com apoio do ambiente moodle: uma abordagem de análise explícita. **Revista Ibérica de Sistemas e Tecnologias de Informação - RISTI**, v. 1, p. 487–500, 2024.

Souza, B. D. de; Moreira, B. G. Detecção de desistência de estudantes em disciplinas ofertadas com apoio do ambiente moodle: uma abordagem de análise implícita. **Anais do Computer on the Beach**, v. 15, p. 064–071, 2024.