

ANÁLISE DE DADOS COM MACHINE LEARNING: Classificação de Alunos em Risco de Evasão Escolar utilizando Modelos de Machine Learning

Matheus Soriano Pimentel

Celso Barreto da Silva

Fábio Fonseca Barbosa Gomes

José Vicente Cardoso Santos

Resumo: Neste artigo, exploramos o uso de modelos de aprendizado de máquina para classificar alunos em risco de evasão escolar, um problema complexo e historicamente relevante no Brasil. O aprendizado de máquina, uma subárea da inteligência artificial, é empregado para identificar os principais atributos que podem levar um aluno a evadir de um curso superior. O aprendizado de máquina é definido como um campo que vai além da inteligência humana, focando não apenas na compreensão, mas também na construção de entidades inteligentes. O problema da evasão escolar é multifacetado, abrangendo questões financeiras, sociais e psicológicas, variando de acordo com o contexto educacional. É destacada a gravidade da evasão escolar no ensino superior e sua influência negativa na qualidade da educação e na sociedade. O objetivo principal do artigo é demonstrar como modelos de aprendizado de máquina podem ser aplicados para prevenir a evasão escolar, com foco na identificação de alunos em risco. Os objetivos específicos incluem a apresentação de diferentes modelos de aprendizado de máquina e o levantamento bibliográfico acerca da utilização destes modelos na análise do problema de evasão. A metodologia aplicada terá como foco a utilização de recursos que estarão disponíveis através do cenário de Mineração de Dados Educacionais, as variáveis que serão analisadas, irão ser recolhidas de diversas bases de dados públicas, pré-processadas e interpretadas através de algoritmos de mineração diferentes para que assim seja possível demonstrar qual dos algoritmos exibidos terão um desempenho melhor em comparação. A pesquisa utiliza técnicas avançadas de análise de dados, incluindo a seleção de atributos relevantes, a criação de modelos de machine learning e a avaliação de métricas de desempenho. Os modelos são alimentados com dados educacionais, como desempenho acadêmico, características socioeconômicas e demográficas. A análise desses dados revela padrões e tendências que indicam quais alunos estão em maior risco de evasão. A importância desse estudo reside na capacidade de instituições de ensino adotarem estratégias proativas para apoiar alunos em risco, oferecendo orientação acadêmica, suporte financeiro ou intervenções psicológicas. A aplicação eficiente de modelos de aprendizado de máquina na prevenção da evasão escolar pode melhorar significativamente as taxas de conclusão dos cursos e mitigar o impacto negativo da evasão na sociedade. Este artigo destaca a relevância do aprendizado de máquina na solução de problemas educacionais complexos, como a evasão escolar. O estudo apresenta uma abordagem promissora para identificar alunos em risco e fornece insights valiosos para aprimorar a qualidade do ensino superior, ressaltando o potencial contínuo do aprendizado de máquina em diversas áreas, incluindo a educação.

Palavras-chave: Machine Learning, Mineração de Dados, Evasão Escolar.

INTRODUÇÃO

Machine Learning (Aprendizado de Máquina) é um ramo da Inteligência Artificial que se define como um campo que vai ainda mais além da inteligência humana, pois ela não tenta apenas compreender, mas também construir entidades inteligentes (RUSSELL, 2013).

O Aprendizado de Máquina é utilizado muitas das vezes com o objetivo de análise de dados de diversos temas, no caso desse texto, será tratado o tema de evasão escolar, mais especificamente na classificação dos alunos que estão em risco de evasão escolar.

O tema da evasão escolar no Brasil sempre foi historicamente complexo pois abrange diversos subtemas que estão agregados ao problema da evasão. Isso inclui problemas financeiros, sociais ou até psicológicos e isso pode variar de um contexto educacional para o outro (LOBO E SILVA et al., 2007).

Segundo HUTAGOL (2019), que aborda evasão escolar, porém utilizando a mineração de dados como um recurso de pesquisa para a prevenção da evasão escolar no ensino superior.

A evasão não é uma novidade, porém é um problema sério que atrai a atenção de pesquisadores devido ao seu impacto na diminuição dos valores do ensino superior e pode ser um impacto adverso no meio social.

A metodologia aplicada terá como foco a utilização de recursos que estarão disponíveis através do cenário de Mineração de Dados Educacionais, as variáveis que serão analisadas, irão ser recolhidas de diversas bases de dados públicas, pré-processadas e interpretadas através de algoritmos de mineração diferentes para que assim seja possível demonstrar qual dos algoritmos exibidos terão um desempenho melhor em comparação.

O objetivo principal deste artigo é apresentar e demonstrar quais são os atributos mais relevantes que possam levar um aluno a evadir um curso superior, seja em uma instituição particular ou pública.

Os objetivos específicos serão focados em realizar um levantamento bibliográfico através de trabalhos correlatos sobre o tema, apresentar alguns modelos de *machine learning* relevantes para o estudo e dar uma explicação

geral sobre quais são os problemas relacionados à evasão escolar e porque o *machine learning* possui valor para este estudo.

REFERENCIAL TEÓRICO

A prevenção da evasão escolar é uma preocupação global na área da educação, com implicações significativas para a qualidade do ensino superior e o desenvolvimento social. Para abordar esse problema complexo, este artigo se baseia em um referencial teórico que abrange diversas áreas de conhecimento, incluindo:

Evasão Escolar no Ensino Superior: A evasão escolar é um desafio crítico nas instituições de ensino superior em todo o mundo. Estudos anteriores têm destacado suas causas e consequências, enfocando fatores financeiros, sociais e psicológicos que contribuem para a evasão de alunos.

No Brasil, e mesmo em outros países, a tendência dos estudos sobre evasão, de um modo geral, é a de orientar-se pela proposta do dimensionamento ou quantificação da evasão (CUNHA et al., 2000).

A evasão por si só é um problema muito árduo de ser resolvido pois abrange diversos fatores complexos o que varia na situação de aluno pra aluno no que diz respeito a problemas sociais, financeiros e até psicológicos (Figura 1) (LOBO E SILVA et al., 2007).

Figura 1. Exemplos de Fatores relacionados à evasão

Fatores internos	Fatores externos
Falta de recursos e de segurança nas escolas;	Falta de transporte local/municipal ao estudante;
O excesso de alunos nas salas de aula;	Vulnerabilidade socioeconômica do estudante;
Falta de qualificação dos professores;	Ausência de ambiente/condições de estudo em casa;
Matrizes curriculares e projetos de curso desatualizados e desalinhados com as necessidades atuais do mercado;	Falta de oportunidades de trabalho na área do curso do discente;

Fonte: Fernando W.B.H Filho et al. 2020

Aprendizado de Máquina e Inteligência Artificial: O aprendizado de máquina é uma subárea da inteligência artificial que se concentra na construção de modelos que podem aprender a partir de dados.

A aplicação de aprendizado de máquina na educação tem ganhado destaque como uma abordagem promissora para identificar alunos em risco de evasão.

Ao recolher os dados que serão abordados e tratá-los através de algoritmos de aprendizagem de máquina, é possível gerar um ou mais modelos que demonstrem a probabilidade de um ou mais eventos acontecerem. Como por exemplo a prevenção de acidentes e a prevenção de tendências (Mueller et al. 2016).

Seguindo este conceito, alguns artigos e trabalhos correlatos já foram construídos abordando este mesmo tema, como é o caso do artigo “Aplicação de Técnicas de Aprendizado de Máquina para Predição de Risco de Evasão Escolar em Instituições Públicas de Ensino Superior no Brasil” (TEODORO et al., 2020).

Que como já descrito no próprio título, se utiliza do aprendizado de máquina para realizar análises nos riscos de evasão escolar no ensino superior público no Brasil, sendo assim uma análise mais geral do problema.

Outro exemplo é o “Predição de Evasão Escolar na Licenciatura em Computação” (JESUS et al., 2021). Neste artigo é feita uma pesquisa de evasão focada apenas na licenciatura dos cursos de Computação, em que são mostradas e descritas algumas variáveis que podem levar o aluno a querer evadir um curso de Computação, como por exemplo, a dificuldade em entender as disciplinas de lógica de programação.

Mineração de Dados Educacionais: A mineração de dados educacionais envolve a coleta, pré-processamento e análise de dados relacionados ao desempenho acadêmico, características dos alunos e outros atributos relevantes. Essa abordagem é fundamental para extrair insights e padrões que podem auxiliar na identificação de alunos em risco.

A coleta de dados é onde tudo começa, a coleta dos dados é descrita como um passo que vêm depois da compreensão do tema envolvido e também

da definição do problema abordado, a coleta dos dados acontece em paralelo a isso (SARKAR et al., 2018).

Tudo que está à nossa volta é considerado um dado ou um material que pode ser interpretado de formas diferentes, pois os dados não possuem um padrão a ser seguido, existem diferenças quanto ao tamanho do arquivo e o seu formato.

Existem alguns formatos de dados que são utilizados atualmente para a coleta de dados. Entre eles estão: CSV, JSON, XML e HTML (SARKAR et al., 2018).

Modelos de Aprendizado de Máquina na Prevenção da Evasão: A literatura acadêmica também fornece uma base teórica para a aplicação de modelos de aprendizado de máquina na prevenção da evasão escolar.

Isso inclui a discussão de diferentes algoritmos de *machine learning*, técnicas de seleção de atributos e métricas de avaliação de desempenho.

Alguns dos algoritmos mais utilizados para a obtenção e interpretação dos dados são: Regressão Linear Simples, Regressão Múltipla, Regressão Logística e Árvores de Decisão (GRUS, 2015).

Também é documentado o uso de algoritmo de Random Forest como no artigo escrito por (SOARES et al. 2020) onde o foco deste artigo é utilizar a Random Forest na identificação de alunos evadidos e não evadidos dentro do contexto acadêmico.

Partindo deste princípio, temos algumas etapas a serem exploradas para que tudo seja feito da maneira correta. O processo descrito como Descoberta de Conhecimento de Dados, que do inglês *significa Knowledge Discovery in Databases* (KDD), possui as seguintes etapas: seleção, pré-processamento, transformação, mineração de dados e avaliação (FAYYAD et al. 1996).

O PROBLEMA DA EVASÃO ESCOLAR NO BRASIL

Analisando os dados de evasão escolar discente no Brasil, observa-se que a evasão por si só é um problema histórico que consiste em muitos fatores incluídos na convivência do aluno no que diz respeito à decisão do aluno em querer evadir o seu curso.

O estudo da evasão no ensino superior público brasileiro, por exemplo, está inserido no cerne de uma discussão que considera o compromisso institucional de uma universidade com as questões e os problemas de seu tempo e as perspectivas de futuro projetadas para sua realidade (GONÇALVES et al., 2019).

Partindo desse raciocínio, existem muitas interações que buscam levar o aluno a concluir o seu curso superior, porém, por outro lado, também existem interações que podem incentivar a evasão desse aluno no curso em questão.

A exemplo disso, temos o curso de Ciência da Computação onde o aluno que decide fazer este curso, normalmente possui a mentalidade de realizar o curso pensando no mercado de trabalho e nas oportunidades que a formação pode oferecer para o indivíduo.

O problema principal é, a partir de um certo período do curso, principalmente os que começam a abordar temas mais complexos como programação (C, Java e derivados), muitos desses alunos que estavam com a mentalidade de construir uma boa formação com base naquela mentalidade inicial acaba indo embora e levando o aluno a evadir o curso.

Como descrito por (JESUS et al., 2021), os problemas principais que levam a isso são justamente a dificuldade que o aluno pode ter nas disciplinas de programação, seja por falta de conhecimento prévio ou até falta de raciocínio lógico-matemático que é bastante abordado nas disciplinas de programação.

A APLICAÇÃO DO APRENDIZADO DE MÁQUINA NA IDENTIFICAÇÃO DE ALUNOS EM RISCO DE EVASÃO

A metodologia de aprendizado de máquina que é aplicada no risco de evasão escolar de alunos, isso tanto para o meio particular quanto para o modelo público de ensino, sempre vai seguir a mesma ordem de aplicação que consiste em:

- Entendimento do problema,
- Compreensão dos dados,
- preparação dos dados,
- modelagem, avaliação e aplicação (JESUS et al., 2021).

A modelagem irá variar muito dependendo do algoritmo que será utilizado para utilizar estes dados.

A exemplo do artigo escrito por (HUTAGOL, 2019), foi utilizado o algoritmo de Decision Tree (Árvore de Decisão), algoritmo esse que é muito popular e conhecimento na área de aprendizado de máquina especialmente quando se fala em classificação de dados.

O algoritmo pode gerar e medir padrões de dados utilizando uma estrutura em árvore e descreve a relação entre variáveis de forma recursiva particionando inputs em duas partes diferentes (HUTAGOL, 2019).

OBJETIVOS DO ESTUDO

O objetivo principal do estudo abordado neste artigo é realizar um levantamento de dados acerca da evasão escolar nos cursos superiores do Centro Universitário Jorge Amado (UNIJORGE) com o intuito de tentar prevenir ao máximo a evasão na instituição.

Visto que, este problema influencia diretamente no desempenho tanto do aluno quanto no da instituição que pode acabar perdendo oportunidades de novos alunos se matricularem ou conhecerem a universidade futuramente.

METODOLOGIA DE MINERAÇÃO DE DADOS EDUCACIONAIS

O processo de mineração de dados educacionais possui alguns passos a serem seguidos para ser aplicado em estudos. Segundo PAL (2012), o processo de mineração começa primeiro com a preparação dos dados, onde neste processo é feita a coleta dos dados que serão estudados e analisados, que no caso deste artigo citado são informações de estudantes do *Institute of Engineering and Technology* da universidade VBS *Purvanchal University* localizada em Jaunpur na Índia por um período de 5 anos.

O segundo passo é descrito como seleção e transformação de dados onde o foco é pegar todos esses dados que foram coletados e classificá-los em variáveis para que assim seja possível uma manipulação e entendimento melhor do que será feito para o passo seguinte que é a implementação do modelo de mineração.

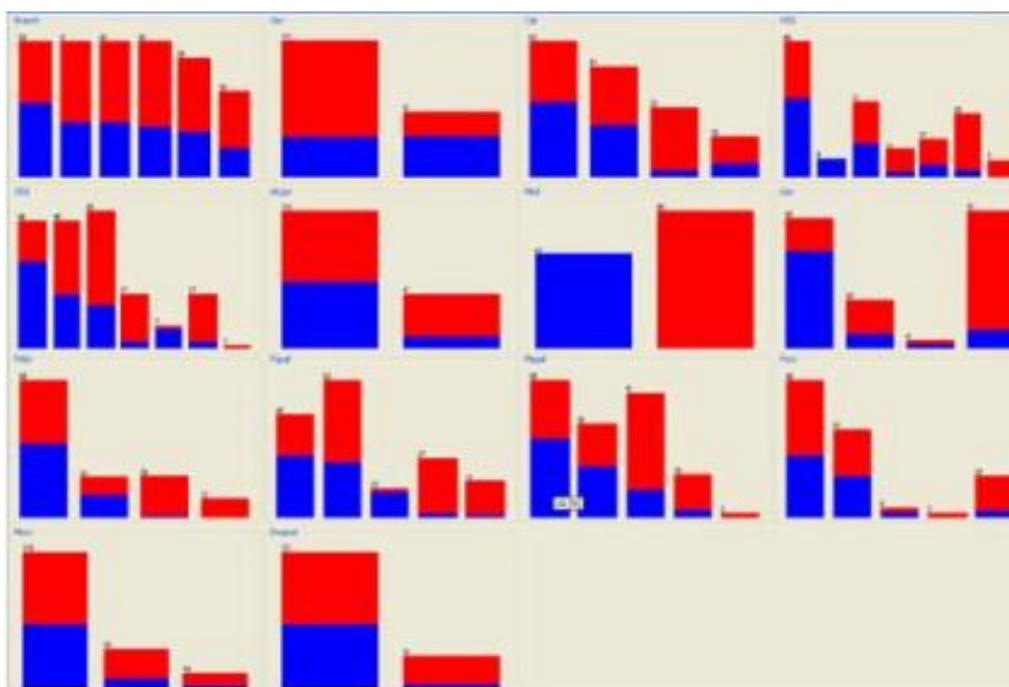
A implementação do modelo de mineração a exemplo dos artigos escritos (PAL, 2012) e (AMORIM et al., 2008), é utilizado a ferramenta *Weka* (WEKA, 2008).

A ferramenta contém uma larga coleção de algoritmos de *machine learning* todos escritos em Java. A ferramenta *Weka* possui algoritmos de regressão, classificação, *clustering*, visualização e pré-processamento de dados.

Para utilizar a ferramenta da maneira correta, é necessário preparar e converter os dados para o formato arff, pois é o formato compatível com a *Weka* (PAL, 2012).

Depois de realizar esses dois passos anteriores, finalmente chega a parte de resultados e discussões sobre a visualização dos dados que é feita através de um gráfico que analisa os dados visualmente e dá uma boa amostra da distribuição de valores que foi obtida através do modelo de mineração utilizado (Figura 2).

Figura 2. Exemplo de gráficos gerados pela ferramenta Weka



Fonte: PAL, 2012.

Análise de Dados e Modelos de Machine Learning

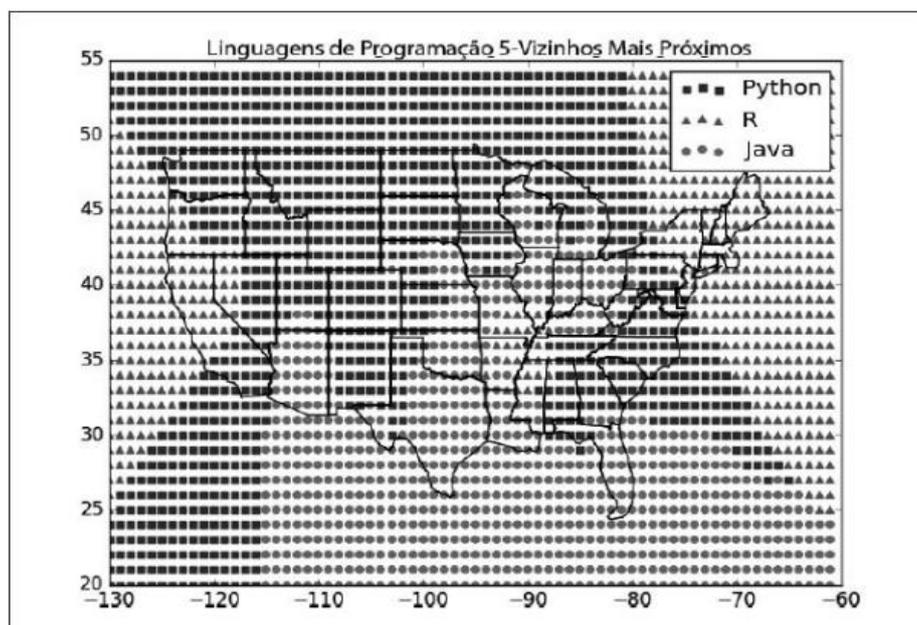
A análise de dados no ramo de Data Science é um caminho que todo cientista de dados tem que seguir para que possa prosseguir com as suas pesquisas.

Porém, como todo cientista de dados, será necessário passar por uma embaraçosa grande fração do seu tempo adquirindo, limpando e transformando dados (GRUS, 2015).

Para trabalhar com estes dados que serão adquiridos ao longo do tempo, existe à disposição diversos modelos de visualização de dados que podem ser usados e manipulados para que a pesquisa realizada seja discutida de forma intuitiva.

Segundo (GRUS, 2015), temos a introdução do primeiro modelo que se denomina como “K-Vizinhos mais próximos”. A descrição deste primeiro modelo diz que ele é um dos modelos preditivos mais simples de se utilizar pois não possui premissas matemáticas e não requer nenhum tipo de maquinário pesado (Figura 3).

Figura 3. Exemplo de K-Vizinhos mais próximos

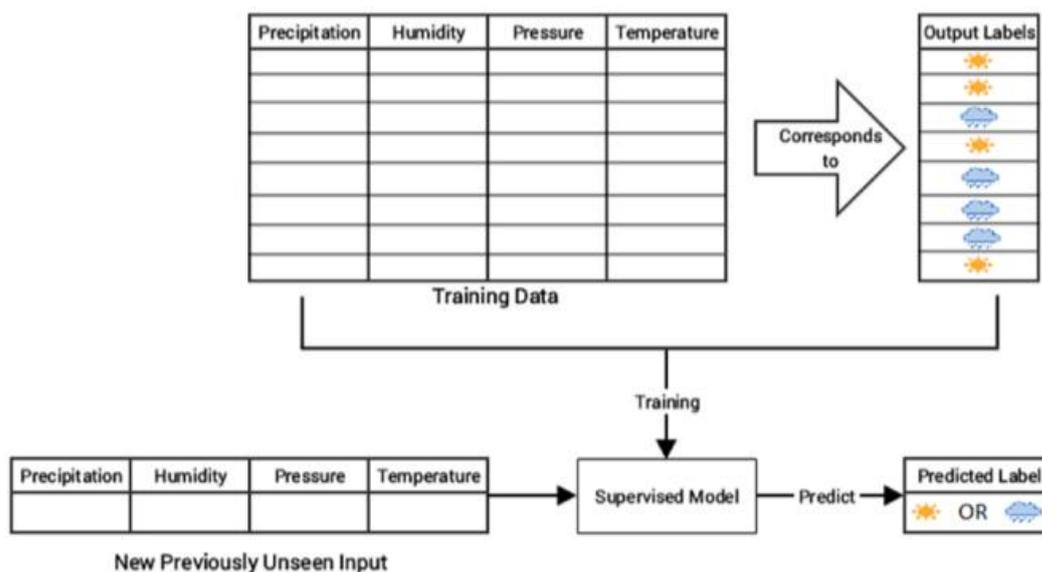


Fonte: GRUS, 2015.

Seguindo agora para algo mais sofisticado, temos os modelos de classificação e regressão que são descritos como modelos de aprendizado supervisionado.

Segundo (SARKAR et al. 2018), o modelo de classificação são “tasks” onde o objetivo chave é realizar uma previsão em tabelas de saída ou respostas que são categóricas de acordo com a natureza dos dados de entrada baseado no que o modelo aprendeu na fase de treinamento e aprendizado (Figura 4).

Figura 4. Como funciona o modelo Classificação

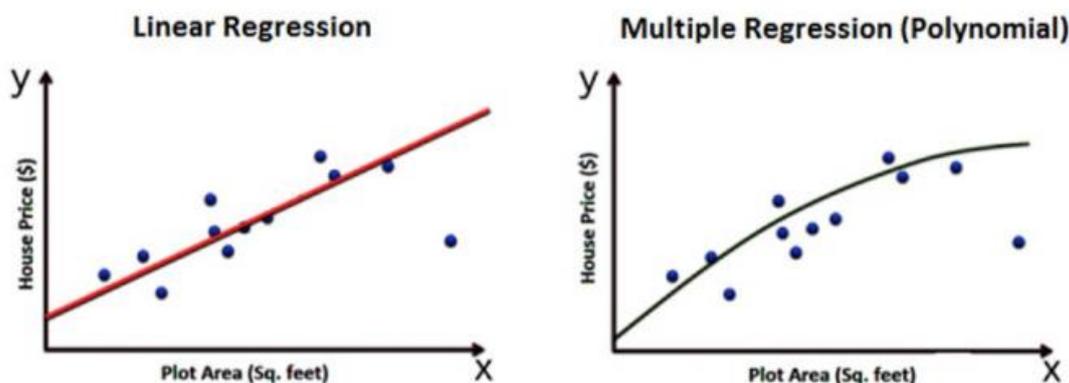


Fonte: SARKAR et al., 2018.

Já o modelo de regressão tem como objetivo realizar uma estimativa de valores que podem ser descritos como “*regression tasks*”.

Métodos de regressão são treinados a partir de amostras de dados de entrada tendo respostas de saída que são contínuos valores numéricos diferente do modelo de classificação que possui discretas categorias e classes (Figura 5).

Figura 5. Regressão Linear e Regressão Múltipla



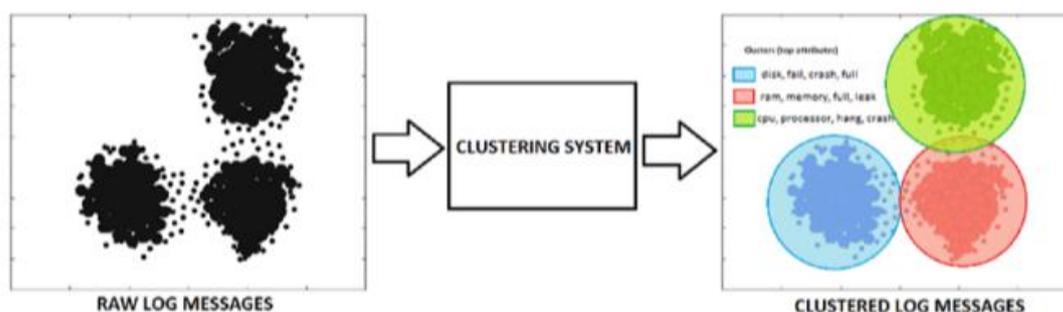
Fonte: SARKAR et al., 2018.

Os próximos modelos citados no mesmo livro, são modelos de aprendizado não-supervisionado. Os modelos em questão são: *Clustering*, redução de dimensionalidade, detecção de anomalia e mineração de regras de associação.

Começando pelo *clustering* segundo (SARKAR et al., 2018), o modelo *clustering* consiste em um método de aprendizado de máquina que tenta encontrar padrões de similaridade e relação a partir de amostras de dados em algum *dataset* e então aglomerá-los em vários grupos.

Sendo que cada um desses grupos aglomerados de amostras de dados possui alguma similaridade, baseado em atributos e recursos inerentes (Figura 6).

Figura 6. Método Clustering

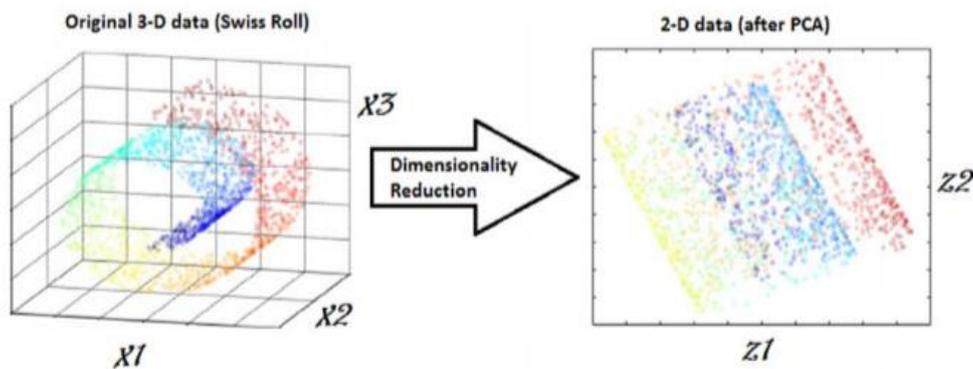


Fonte: SARKAR et al., 2018.

O próximo modelo descrito chama-se redução de dimensionalidade, segundo *Dijanpan Sarkar* (2018) esse método se baseia em reduzir o número de variáveis extraíndo ou selecionando um “set” de principais ou representativos recursos (Figura 7).

Existem diversos algoritmos de redução de dimensionalidade sendo os três mais populares: *Principal Component Analysis (PCA)*, *nearest neighbors*, e *discriminant analysis*.

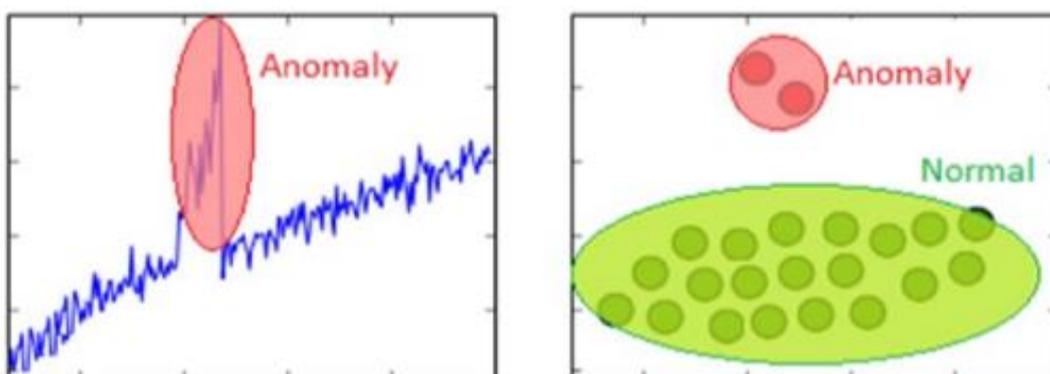
Figura 7. Redução de Dimensionalidade



Fonte: SARKAR et al., 2018.

A detecção de anomalia é um processo que também pode ser chamado de detecção de valores discrepantes, onde o objetivo é encontrar ocorrências de raros eventos ou observações que tipicamente não ocorrem normalmente baseado em amostras de dados históricos (Figura 8).

Figura 8. Detecção de Anomalia

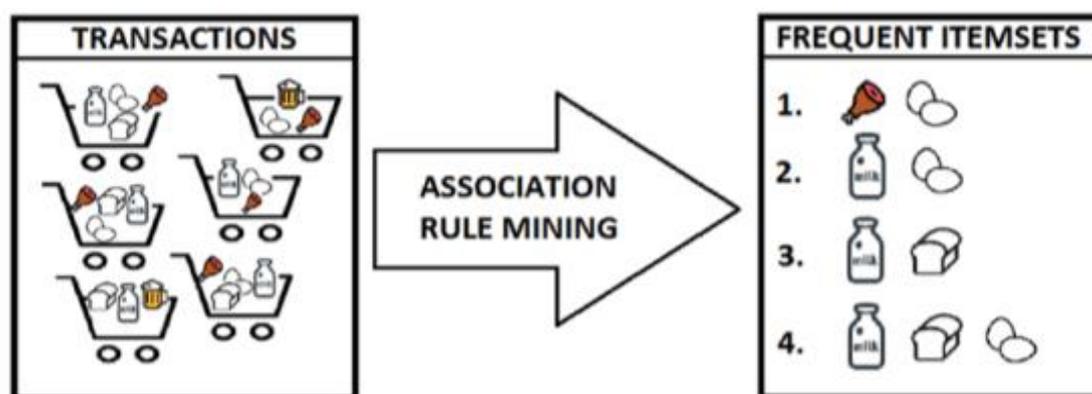


Fonte: SARKAR et al., 2018.

Por fim, temos a mineração de regras de associação que basicamente é um método de mineração que é usado para examinar e analisar largos *datasets* transacionais a fim de encontrar padrões e regras de interesse.

Esses padrões representam interessantes relações e associações através de vários itens a partir de transações (Figura 9).

Figura 9. Mineração de Regras de Associação



Fonte: SARKAR et al., 2018.

Importância da Prevenção da Evasão Escolar

Segundo (Lobo e Silva et al., 2007), a evasão acarreta na perda de estudantes que iniciam, porém não terminam seus cursos. Essa perda de alunos leva a desperdícios sociais, econômicos e acadêmicos no setor privado.

Já no setor público, leva a gastos públicos investidos sem o devido retorno. Em ambos os casos, a evasão é uma fonte de ociosidade de professores, funcionários, equipamentos e espaço físico (Lobo e Silva et al., 2007).

Partindo destes conceitos, a importância de prevenir a evasão escolar no setor privado de ensino é principalmente pelo fato de que a educação no setor privado é um importante gerador de receitas.

Porém, enquanto é despendido cerca de 2% a 6% das receitas das instituições em marketing para atrair novos alunos, o mesmo não parece ser feito para manter os estudantes já matriculados (Lobo e Silva et al., 2007).

Já no setor público, como é algo financiado pelo governo diretamente, a importância de prevenir a evasão neste setor é para evitar gastos desnecessários de dinheiro público que vai ser investido na educação. Até porque não faz sentido investir em uma instituição pública que está sofrendo constantemente com problemas de evasão.

CONCLUSÃO

A mineração de dados e análise de dados feita com o auxílio de *machine learning* possui um potencial enorme para ser explorado em diversas áreas diferentes da tecnologia.

É fato de que, como a Inteligência Artificial tecnicamente falando é algo bem novo e tende a ser explorado com bastante cautela, ainda gera bastante confusão nas pessoas que se interessam e querem aprender sobre o tema.

Isso é questão de tempo até acontecer uma dominância plena sobre o que é a Inteligência Artificial e o que ela pode oferecer de benefícios para a humanidade como um todo.

Foi realizado durante a construção deste artigo científico um levantamento acerca das variáveis e atributos mais relevantes no tema da evasão escolar. O objetivo geral foi alcançado, pois foram apresentadas variáveis externas e internas, além de descrever essas variáveis através do ambiente público e privado de ensino.

Para que o objetivo geral fosse alcançado, foi necessário o delineamento de atividades específicas como, um levantamento bibliográfico através de pesquisas de trabalhos correlatos acerca do tema de evasão escolar, onde objetivava-se a explicar sobre o problema de evasão escolar e sobre o aprendizado de máquina focado neste problema.

Além disso, foram apresentados alguns algoritmos e modelos de aprendizado de máquina que foram utilizados para realizar análises e pesquisas sobre o tema. Todos os objetivos foram concluídos através de uma apresentação de imagens e descrições construídas ao longo do desenvolvimento do artigo.

Sobre a Evasão Escolar, também é um fato de que esse é um problema muito recorrente em diversas instituições ao redor do mundo e não apenas no Brasil. É muito importante que haja formas de amenizar esse problema da maneira mais eficiente possível e um dos caminhos que podem levar a solução será com a ajuda do Aprendizado de Máquina.

O Aprendizado de Máquina pode muito bem começar a guiar de forma bem eficiente sobre o que pode ser feito acerca do problema. A coleta de dados é realizada com a finalidade de utilizar o Aprendizado de Máquina para que sirvam como ferramenta para o processo decisório de instituições de ensino e assim a realização de descoberta de informações sobre a causa da evasão.

É fato que isso ainda é um desafio, porém, a intenção de apresentar, demonstrar e conscientizar a quem ainda não conhece sobre o tema ou conhece muito pouco sobre, é bastante importante para que haja discussões e é com discussões que se pode chegar no objetivo final que é a prevenção da Evasão Escolar.

Perspectivas Futuras

Acredita-se que com a construção deste artigo, seja possível levar para o máximo de pessoas como funciona a Mineração e Análise de Dados e também quais fatores que podem levar a Evasão Escolar de um aluno.

Sendo assim, acredita-se também que este artigo possa servir de auxílio a instituições de ensino que procuram entender melhor sobre o tema de Evasão e o que elas podem fazer para que seja possível prevenir esse problema da melhor maneira possível.

Referências

AMORIM, Maurício et al. **Técnicas de Aprendizado de Máquina Aplicadas na Previsão de Evasão Acadêmica**. 2008. Artigo Científico – XIX Simpósio Brasileiro de Informática na Educação (SBIE 2008). Disponível em: [\(PDF\) Técnicas de Aprendizado de Máquina Aplicadas na Previsão de Evasão Acadêmica \(researchgate.net\)](#). Acesso em: 12/10/2023.

CUNHA, Aparecida et al. **Evasão do Curso de Química da Universidade de Brasília: A Interpretação do Aluno Evadido**. 2000. Artigo Científico – Química Nova Educação. Disponível em: [\(PDF\) Evasão do curso de química da Universidade de Brasília: a interpretação do aluno evadido \(researchgate.net\)](#). Acesso em: 15/10/2023.

FAYYAD, Usama. **Advances in Knowledge Discovery & Data Mining**. Cambridge, Massachusetts: MIT Press, 1996. 626 p.

FILHO, Fernando et al. **Predição de Evasão Utilizando Técnicas de Classificação: Um Estudo de Caso do Instituto Federal do Ceará**. 2020. Artigo Científico – Instituto Federal do Ceará (IFCE). Disponível em: [Vista do Predição de Evasão Utilizando Técnicas de Classificação: Um Estudo de Caso do Instituto Federal do Ceará \(sbc.org.br\)](#). Acesso em: 15/10/2023.

GONÇALVES, Oldair. **Mineração de Dados e Evasão Estudantil: Analisando o Curso de Nível Superior do IFES**. 2019. Artigo Científico – XII CASI – Congresso de Administração, Sociedade e Inovação. Disponível: [\(PDF\) Mineração de dados e Evasão estudantil: Analisando o curso de nível superior do Ifes \(researchgate.net\)](#). Acesso em: 15/10/2023.

GRUS, Joel. **Data Science do Zero**. Rio de Janeiro: Alta Books, 2016. 336 p.

HUTAGOL, Nindhia. **Predictive Modelling of Student Dropout Using Ensemble Classifier Method in Higher Education**. 2019. Artigo Científico – ASTES Journal. Disponível em: [Predictive Modelling of Student Dropout Using Ensemble Classifier Method in Higher Education - Advances in Science, Technology and Engineering Systems Journal \(astesj.com\)](#). Acesso em: 15/10/2023.

JESUS, Hiago et al. **Predição de Evasão Escolar na Licenciatura em Computação**. 2021. Artigo Científico – Revista Brasileira de Informática na Educação – RBIE. Disponível em: [\(PDF\) Predição de Evasão Escolar na Licenciatura em Computação \(researchgate.net\)](#). Acesso em: 17/10/2023.

LOBO E SILVA et al. **A Evasão no Ensino Superior Brasileiro**. 2007. Artigo Científico – Instituto Lobo para o Desenvolvimento da Educação, da Ciência e da Tecnologia. Disponível em: [\(PDF\) A evasão no ensino superior brasileiro \(researchgate.net\)](#). Acesso em: 20/10/2023.

MUELLER, John et al. **Machine Learning for Dummies**. Estados Unidos: For Dummies, 2016. 464 p.

PAL, Saurabh. **Mining Educational Data to Reduce Dropout Rates of Engineering Students**. 2012. Artigo Científico – International Journal of Information Engineering and Electronic Business. Disponível em: [\(PDF\) Mining Educational Data to Reduce Dropout Rates of Engineering Students \(researchgate.net\)](#). Acesso em: 17/10/2023.

RUSSELL, Stuart. **Inteligência Artificial**. Rio de Janeiro: Elsevier, 2013. 1016 p.

SARKAR, Dipanjan. **Practical Machine Learning with Python: A Problem-Solver's Guide to Building Real-World Intelligent Systems**. Bangalore, Índia: Apress, 2018. 555 p.

SOARES, Leandra et al. **Aplicação de Técnicas de Aprendizado de Máquina em um Contexto Acadêmico Com Foco na Identificação de Alunos Evadidos e Não Evadidos**. 2020. Artigo – Humanidades e Inovação: Educação formal e não formal, cultura e currículo III.

TEODORO, Leonardo et al. **Aplicação de Técnicas de Aprendizado de Máquina para Predição de Risco de Evasão Escolar em Instituições Públicas de Ensino Superior no Brasil**. 2020. Artigo Científico – Revista Brasileira de Informática na Educação – RBIE. Disponível em: [Aplicação de Técnicas de Aprendizado de Máquina para Predição de Risco de Evasão Escolar em Instituições Públicas de Ensino Superior no Brasil | Teodoro | Revista Brasileira de Informática na Educação \(usp.br\)](#). Acesso em: 23/10/2023.