



UNIVERSIDADE FEDERAL FLUMINENSE
INSTITUTO DE COMPUTAÇÃO
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

JORGE AUGUSTO VIEIRA GOMES DA COSTA CARVALHO
NICOLE LUONGO AGUIAR

**O PAPEL DO BIG DATA NA EDUCAÇÃO: APRIMORANDO MÉTODOS DE
ENSINO E ENRIQUECENDO O PROCESSO DE APRENDIZAGEM**

Niterói
2023

JORGE AUGUSTO VIEIRA GOMES DA COSTA CARVALHO
NICOLE LUONGO AGUIAR

**O PAPEL DO BIG DATA NA EDUCAÇÃO: APRIMORANDO MÉTODOS DE
ENSINO E ENRIQUECENDO O PROCESSO DE APRENDIZAGEM**

Trabalho de conclusão de curso apresentado
ao curso de Bacharelado em Sistemas de
Informação, como requisito parcial para
conclusão do curso.

Orientador:
Prof. Dr. Esteban Gonzalez Clua

Niterói
2023

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

A282p Aguiar, Nicole Luongo
O PAPEL DO BIG DATA NA EDUCAÇÃO: APRIMORANDO MÉTODOS DE
ENSINO E ENRIQUECENDO O PROCESSO DE APRENDIZAGEM / Nicole
Luongo Aguiar, Jorge Augusto Vieira Gomes da Costa Carvalho. -
2023.
41 f.: il.

Orientador: Esteban Walter Gonzalez Clua.
Trabalho de Conclusão de Curso (graduação)-Universidade
Federal Fluminense, Instituto de Computação, Niterói, 2023.

1. Tecnologia Educacional. 2. Acurácia dos Dados. 3.
Método de ensino. 4. Produção intelectual. I. Carvalho,
Jorge Augusto Vieira Gomes da Costa. II. Clua, Esteban Walter
Gonzalez, orientador. III. Universidade Federal Fluminense.
Instituto de Computação.IV. Título.

CDD - XXX

JORGE AUGUSTO VIEIRA GOMES DA COSTA CARVALHO
NICOLE LUONGO AGUIAR

**O PAPEL DO BIG DATA NA EDUCAÇÃO: APRIMORANDO MÉTODOS DE
ENSINO E ENRIQUECENDO O PROCESSO DE APRENDIZAGEM**

Trabalho de conclusão de curso
apresentado ao curso de Bacharelado em
Sistemas de Informação, como requisito
parcial para conclusão do curso.

Aprovada em 07 de dezembro de 2023.

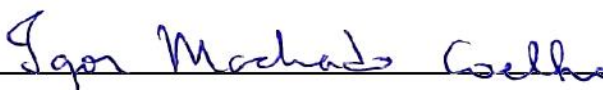
BANCA EXAMINADORA



Prof. Dr. Esteban Gonzalez Clua (Orientador) - UFF



Prof.^a D.Sc Luciana Castro Salgado - UFF



Prof. Dr. Igor Machado Coelho - UFF

Niterói
2023

AGRADECIMENTOS

Gostaríamos de agradecer a todos os nossos familiares e amigos que sempre estiveram ao nosso lado, nos apoiando ao longo de nossa jornada acadêmica.

Aos nossos professores pelo conhecimento transmitido e, ao nosso orientador, Esteban, obrigado por nos guiar durante os meses de desenvolvimento deste trabalho.

RESUMO

O advento da era digital trouxe uma vasta quantidade de dados, inaugurando a era do Big Data. Essa tendência, que vai além dos domínios da ciência de dados e da tecnologia da informação, tem alcançado diversos setores, incluindo a educação. Este estudo propõe-se a explorar a influência do Big Data na educação, enfocando como ele está transformando métodos ultrapassados e ineficazes de ensino e enriquecendo o processo de aprendizagem. A integração do Big Data no setor educacional representa uma mudança paradigmática, trazendo consigo novos desafios e abrindo um leque de possibilidades para professores, alunos e instituições educacionais. A análise em questão, fundamentada na premissa de que o Big Data tem o potencial de aprimorar significativamente a qualidade do ensino e do aprendizado, visa compreender como esta tecnologia pode ser efetivamente aplicada no contexto educacional. Exploramos a capacidade do Big Data de processar e analisar grandes volumes de dados em tempo real, o que nos permite entender melhor as necessidades e comportamentos dos alunos. Isso, por sua vez, possibilita o desenvolvimento de métodos de ensino mais personalizados e adaptativos, que promovem um ambiente educacional mais dinâmico e interativo. Analisamos também as implicações práticas da aplicação do Big Data na educação, incluindo estratégias de coleta de dados, as tecnologias envolvidas e os desafios éticos no uso de informações sensíveis dos alunos. Através de exemplos reais, investigamos como o Big Data está sendo implementado para melhorar o ensino e o aprendizado, tanto em escolas quanto em plataformas de ensino digital, proporcionando percepções enriquecedoras para educadores.

Palavras-chave: Big Data. Educação. Coleta de Dados. Métodos de Ensino.

ABSTRACT

The advent of the digital age has brought vast amounts of data, ushering in the era of Big Data. This trend, which goes beyond the realms of data science and information technology, has reached various sectors, including education. This study sets out to explore the influence of Big Data on education, focusing on how it is transforming outdated and ineffective teaching methods and enriching the learning process. The integration of Big Data into the education sector represents a paradigm shift, bringing with it new challenges and opening up a range of possibilities for teachers, students and educational institutions. The analysis, based on the premise that Big Data has the potential to significantly improve the quality of teaching and learning, aims to understand how this technology can be effectively applied in the educational context. We explore Big Data's ability to process and analyze large volumes of data in real time, which allows us to better understand student's needs and behaviors. This, in turn, makes it possible to develop more custom and adaptive teaching methods, which promote a more dynamic and interactive educational environment. We also look at the practical implications of applying Big Data in education, including data collection strategies, the technologies involved and the ethical challenges in using sensitive student information. Through real examples, we investigate how Big Data is being implemented to improve teaching and learning, both in schools and on digital learning platforms, providing enriching insights for educators.

Keywords: Big Data. Education. Data Collect. Learning Methods.

SUMÁRIO

1	INTRODUÇÃO	07
1.1	Objetivos	07
1.2	Metodologia	08
1.3	Organização do trabalho	09
2	O QUE É O BIG DATA	10
3	APLICAÇÕES DO BIG DATA NA EDUCAÇÃO	12
3.1	Análise preditiva de desempenho acadêmico	12
3.2	Prevenção de evasão escolar e retenção de estudantes	13
3.3	Personalização do aprendizado	15
3.4	Previsão de demandas futuras e alocação eficiente de recursos	16
3.5	Aprimoramento contínuo do ensino	17
4	EXPERIÊNCIAS BEM SUCEDIDAS	19
4.1	Duolingo	19
4.2	Summit Public Schools	22
4.3	Khan Academy	24
4.4	Purdue University	27
4.5	Arizona State University	29
5	DESAFIOS DO BIG DATA	31
6	CONSIDERAÇÕES FINAIS	33
	REFERÊNCIAS	35

1 INTRODUÇÃO

Com o desenvolvimento da tecnologia e a ascensão da internet, a quantidade de dados gerados a cada minuto pelos mais diversos tipos de aparelhos tecnológicos aumenta significativamente, resultando no surgimento do conceito de Big Data. “Big” trazendo a ideia de grande e “Data” de dados, fazendo referência a um grande volume de dados. Este conceito, que vem crescendo muito nos últimos anos (SAGIROGLU; SINANC, 2013), possui o intuito de estudar maneiras de tratar e analisar esses dados em larga escala que não conseguem ser trabalhados em sistemas tradicionais, direcionando para ações mais estratégicas e assertivas (SABHARWAL; MIAH, 2021). Dessa maneira, a revolução digital tem transformado significativamente vários setores, e a educação não é exceção. A capacidade de extrair e analisar informações relevantes desse enorme conjunto de dados não apenas muda a maneira com a qual compreendemos a educação, mas também possibilita descobertas que podem auxiliar no aprimoramento do ensino.

Nos dias atuais, a sociedade exige indivíduos altamente capacitados, aptos a enfrentar desafios complexos e em constante evolução ao mesmo tempo que, embora muitos avanços tenham ocorrido na educação, alguns métodos ultrapassados ainda persistem, impactando a eficácia do ensino (SANTOS, 2016). Tal revolução tem impactado diversos setores, incluindo a educação. Nesse contexto, a integração do Big Data na educação não se apresenta apenas como uma escolha, mas como uma necessidade para garantir que o processo educacional esteja alinhado com as demandas do século XXI. Ao examinar de perto as implicações e os benefícios dessa interseção entre tecnologia e educação, podemos não apenas compreender as potencialidades do Big Data, mas também delinear como sua aplicação estratégica pode moldar o futuro do ensino.

1.1. Objetivos

Este trabalho tem como objetivo demonstrar e analisar de que formas o Big Data pode ser uma alternativa para aprimorar métodos de ensino ultrapassados para enriquecer o processo de aprendizado, diminuir taxas de evasão escolar e

Entre os objetivos específicos estão:

- Entender de que formas o Big Data pode aprimorar diversas questões relacionadas à educação.
- Analisar exemplos reais de ferramentas, aplicativos e instituições que se utilizaram do big data para esse aprimoramento
- Discutir as dificuldades que podem ser encontradas implementando essa técnica.

1.2. Metodologia

O presente estudo adota uma abordagem descritiva qualitativa que é justificada pela natureza teórica da pesquisa, que se baseia em fontes como artigos acadêmicos, documentos educacionais e referências bibliográficas relacionadas ao tópico abordado neste trabalho.

O estudo concentra-se no papel do big data na educação: analisando métodos de ensino e enriquecendo o processo de aprendizagem, utilizando métodos qualitativos que permitiram uma análise abrangente das práticas de ensino e das percepções sobre o tema em questão. A seguir está a organização das etapas para a metodologia:

- Etapa 1: Busca por artigos e referências que possuem conteúdos relevantes e relacionados ao tema discutido.
- Etapa 2: Análise dos documentos, revisando a bibliografia identificada, com foco na extração de informações chave.
- Etapa 3: Pesquisa e análise de exemplos concretos da aplicação do big data em diferentes contextos educacionais, incluindo instituições públicas e privadas, plataformas digitais e aplicativos móveis.
- Etapa 4: Seleção de 5 exemplos dentre os coletados, levando em consideração as diferentes categorias educacionais buscadas e a identificação de estudos e entrevistas que fornecessem dados relevantes sobre a implementação tecnológica discutida.

- Etapa 5: Identificação e análise de desafios tanto estruturais quanto éticos relacionados ao uso do Big Data.

1.3. Organização do trabalho

No capítulo 2, é descrita a fundamentação teórica do trabalho, apresentando conceitos sobre o big data e outras tecnologias correlatas.

No capítulo 3, são discutidos as aplicações do big data e suas tecnologias no contexto educacional e suas implicações positivas.

No capítulo 4, são apresentados exemplos práticos que fazem alusão ao tema da pesquisa.

No capítulo 5, são discutidos os desafios e dificuldades da implementação do big data na área da educação.

No capítulo 6, são realizadas as considerações finais, sintetizando as principais descobertas e conclusões do estudo.

2 O QUE É O BIG DATA?

No âmbito da Tecnologia da Informação, o termo "Big Data" designa um amplo espectro de técnicas, instrumentos e metodologias elaboradas com o propósito de processar e analisar extensas quantidades de dados, englobando tanto dados estruturados quanto aqueles que não possuem estrutura definida. Essa metodologia supera as técnicas tradicionais de coleta, armazenamento e análise de dados, objetivando a extração de informações relevantes de conjuntos de dados amplos, diversificados e em constante transformação (SABHARWAL; MIAH, 2021).

Ademais, as características singulares do Big Data são frequentemente resumidas nos chamados "5Vs": Volume, que se refere à imensa quantidade de dados gerados, coletados e processado; Velocidade, relacionada à rapidez com que os dados são produzidos, analisados e transferidos entre diferentes sistemas e dispositivos; Variedade, que contempla a diversidade das fontes de dados; Veracidade, concentrando-se na qualidade e exatidão dos dados; e Valor, referindo-se a diferentes tipos de benefícios que podem ser obtidos do processamento e análise das informações (YOUNAS, 2019).

Do ponto de vista técnico, o Big Data envolve a utilização de tecnologias avançadas para administrar a complexidade e o volume desses dados. Entre elas podemos citar:

- **Computação em nuvem:** Esta tecnologia permite o armazenamento e processamento de dados em servidores remotos, acessíveis via internet. A computação em nuvem oferece flexibilidade, escalabilidade e eficiência, possibilitando que as organizações gerenciem grandes volumes de dados sem a necessidade de infraestrutura física própria (BERISHA, B.; MĚZIU, E.; SHABANI, I., 2022). Este modelo de computação é essencial para o Big Data, pois fornece os recursos necessários para processar e analisar grandes conjuntos de dados de forma eficiente
- **Data mining (Mineração de dados) com algoritmos de aprendizado de máquina (Machine learning):** O Data mining consiste no processo de o processo de descobrir padrões, correlações, tendências e anomalias em grandes conjuntos de dados (HAN, J.; KAMBER, M.; PEI, J., 2011)

se utilizando de algoritmos de aprendizado de máquina, que permitem que os sistemas aprendam e melhorem a partir da experiência sem serem explicitamente programados (MAHESH, 2019). No contexto do Big Data, a mineração de dados em conjunto com a tecnologia do aprendizado de máquina é utilizada para analisar grandes volumes de dados, identificar padrões e fazer previsões.

- **Sistemas distribuídos:** Um sistema distribuído é aquele em que componentes localizados em computadores interligados por uma rede comunicam-se e coordenam suas ações exclusivamente por meio do envio de mensagens (COULOURIS *et al.*, 2005). Esses sistemas são desenhados para dividir e gerenciar o processamento de dados em larga escala entre múltiplos computadores interconectados, denominados 'nós' (DUCANGE, P.; FAZZOLARI, M.; MARCELLONI, F., 2020). Esta configuração não apenas eleva a eficiência no processamento dos dados, mas também acelera significativamente a análise, aproveitando ao máximo a capacidade computacional disponível. Essa abordagem distribuída é essencial para lidar com a magnitude e a complexidade dos desafios apresentados pelo Big Data.

O Big Data desempenha um papel fundamental em diversos campos, incluindo o setor empresarial, a ciência, a saúde e o governo. A capacidade de analisar grandes volumes de dados possibilita a identificação de padrões, tendências e correlações que seriam difíceis de detectar utilizando métodos de análise de dados tradicionais (HARIRI, R.; FREDERICKS, E.; BOWERS, K., 2019). Essa prática tem o potencial de fornecer insights significativos, fomentar a inovação e aumentar a eficiência operacional em variados contextos. Dessa forma, o Big Data representa uma evolução significativa no paradigma de gestão de dados, habilitando organizações e pesquisadores a extrair conhecimentos aprofundados de vastos repositórios de informações, o que contribui para uma tomada de decisão mais informada e eficiente.

3 APLICAÇÕES DO BIG DATA NA EDUCAÇÃO

A integração do Big Data no setor educacional é um tema de crescente relevância em meio às tendências emergentes na educação (WAGNER; ICE, 2012). Essa tecnologia, quando aplicada de maneira eficaz, tem o potencial de revolucionar os paradigmas de ensino e aprendizagem.

No contexto educacional, as informações provenientes do grande volume de dados coletado e analisados, têm 5 tipos (FERREIRA, 2013):

- Dados de identificação do usuário: dados que informam quem são os usuários, as permissões concedidas a eles e informações de cunho demográfico e socioeconômico.
- Dados da interação do usuário com o ambiente de aprendizagem: a experiência que o usuário possui utilizando a plataforma de aprendizagem, quantidade de visitas nas páginas, taxas de cliques e outras diversas métricas.
- Dados sobre conteúdo: se referem a dados que apontam níveis de desempenho, realizando paralelos entre conteúdo escolar e eficácia na aprendizagem.
- Dados relacionados aos sistemas de informação das organizações educacionais: dados que já são recorrentes e coletados pelas escolas como notas e históricos disciplinares.
- Dados inferidos sobre os estudantes: referem-se ao comportamento de cada estudante durante o processo de aprendizagem de um conteúdo. A coleta e análise de dados desse tipo envolve uma dificuldade maior, em razão da ampla gama de situações que o estudante pode vir a enfrentar durante o aprendizado.

Utilizando esses dados, são inúmeras as possibilidades de aprimoramento na educação e, nos tópicos abaixo, falaremos sobre as mais relevantes até o momento.

3.1 Análise preditiva de desempenho acadêmico

O Big Data tem emergido como uma ferramenta fundamental na análise preditiva do desempenho acadêmico dos alunos. A aplicação de algoritmos em extensos conjuntos de dados permite a identificação de padrões que predizem o futuro desempenho dos alunos em disciplinas específicas. Modelos de aprendizado profundo, por exemplo, ao serem aplicados aos dados de sistemas de gestão de aprendizado são capazes de realizar tais previsões com precisão (WAHEED *et al.*, 2020). Complementando essa visão, enfatiza-se a importância da avaliação de métodos de classificação e registros de aprendizagem para a predição do desempenho acadêmico, evidenciando que padrões de acesso e ações de navegação nos sistemas de gestão de aprendizagem fornecem informações sobre os estudantes em risco. Esta abordagem permite que as instituições educacionais desenvolvam estratégias corretivas e pedagógicas mais efetivas, possibilitando intervenções antecipadas, como por exemplo, tutorias personalizadas e programas de reforço vitais para a melhoria do desempenho acadêmico.

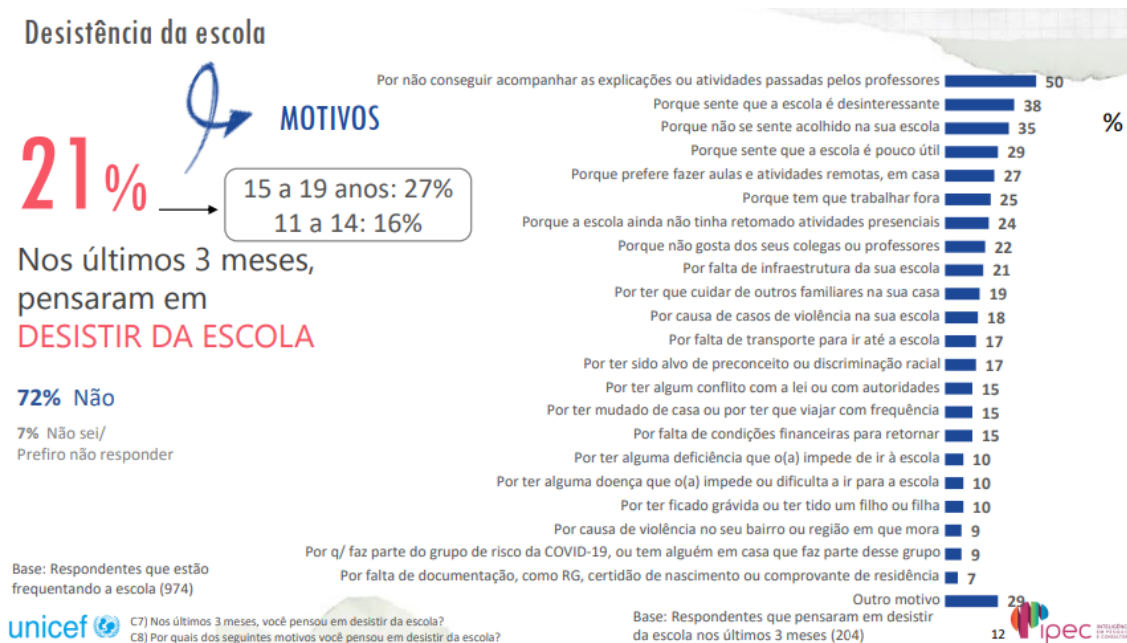
A capacidade de prever tendências futuras na demanda por cursos ou áreas de estudo torna-se uma contribuição valiosa. Por exemplo, uma pesquisa feita em uma universidade da Turquia revelou que as previsões feitas usando apenas três tipos de parâmetros: notas de exames intermediários, dados do departamento e da faculdade, usados como base de dados para um modelo de algoritmos baseado em machine learning, alcançaram uma precisão de classificação de 70-75% (YAĞCI, 2022).

Em resumo, a análise preditiva auxiliada pelo Big Data na educação não apenas antecipa o desempenho dos alunos, mas também fomenta uma abordagem educacional mais proativa e personalizada, possibilitando que as instituições educacionais melhorem constantemente suas práticas de ensino, estabelecendo um ciclo de feedback essencial para a melhoria da qualidade educacional. As informações coletadas por tais técnicas de predição contribuem significativamente para a promoção do sucesso acadêmico e o aprimoramento do processo educacional como um todo.

3.2 Prevenção de evasão escolar e retenção de estudantes

Um estudo realizado pela UNICEF intitulado “Educação brasileira em 2022 – a voz de adolescentes” forneceu diversos dados relacionados à educação. Um desses dados envolve a evasão escolar conforme mostra a figura 1.

Figura 1 - Motivos de desistência escolar



Fonte: UNICEF, 2022

Analisando os dados fornecidos, percebemos que mais de 20% dos alunos já pensaram em desistir da escola por diversos motivos. A identificação precoce desses sinais de risco permite que as instituições educacionais implementem ações específicas para retenção de estudantes, como programas de tutoria individualizada e apoio emocional. Estudos empíricos destacam que alunos em risco de evasão escolar geralmente apresentam as seguintes condições: reprovação em disciplinas principais, absenteísmo excessivo, falha em avançar para a próxima série, desengajamento na sala de aula, problemas familiares ou condições socioeconômicas ruins (ALVES, C.; LAVOR, L.; PEREIRA, H., 2017). Todos esses fatores contextuais importantes são levados em consideração para compreensão e abordagem de tais riscos, que por sua vez são provenientes de análises do Big Data. Além disso, o desenvolvimento de ferramentas de engajamento estudantil, como plataformas online interativas e aplicativos móveis, também se beneficia do

Big Data, ao permitir monitorar a participação dos estudantes e identificar padrões comportamentais indicativos de motivação e satisfação.

A utilização contínua do Big Data na prevenção da evasão escolar estabelece um ciclo de feedback ininterrupto. A avaliação da eficácia das intervenções implementadas e o ajuste de estratégias com base nos resultados possibilitam que as instituições aprimorem constantemente suas abordagens de retenção de estudantes.

Em síntese, a aplicação do Big Data na prevenção da evasão escolar não se limita à identificação de sinais de risco, mas também oferece ferramentas e estratégias personalizadas para a permanência dos estudantes nos cursos (LEE; CHUNG, 2019). Ao compreender e abordar fatores acadêmicos e contextuais, o Big Data contribui para a criação de ambientes educacionais mais inclusivos e adaptados às necessidades individuais dos estudantes, resultando em impactos positivos na retenção e no sucesso acadêmico.

3.3 Personalização do aprendizado

A personalização do aprendizado, é uma abordagem educacional que se adapta às realidades educacionais específicas de cada aluno. Esta metodologia foca na individualidade do estudante, levando em consideração suas necessidades, interesses e habilidades únicas (ESTEVAM, 2021).

De acordo com a mesma pesquisa da UNICEF apresentada no tópico acima na figura 1, é possível observar que o principal motivo de desistência escolar em 2022 foi "Por não conseguir acompanhar as explicações ou atividades passadas pelo professor". Para contornar esse problema, utilizando-se de análises de Big Data, é possível desenvolver rotas de aprendizado adaptativas, as quais personalizam o percurso educacional. Isso permite que cada aluno avance de acordo com seu próprio ritmo e nível de compreensão. As atividades e desafios são ajustados à medida que o estudante progride, assegurando um ambiente desafiador, mas igualmente acessível.

A análise permite ainda a identificação das áreas onde os alunos enfrentam mais dificuldades. Com base nesses dados, o sistema educacional pode oferecer

materiais didáticos personalizados, como: recomendações de leitura, vídeos educativos, exercícios adaptativos e outras ferramentas de aprendizagem. Estes recursos são adaptados de acordo com o estilo de aprendizagem de cada aluno, otimizando a compreensão e o interesse pelo conteúdo apresentado. Esses pontos fracos são discernidos por meio de análises preditivas, já discutidas anteriormente, possibilitando intervenções específicas, como tutorias personalizadas e revisões de conceitos. Além disso, a coleta contínua de dados, capacita os sistemas educacionais a fornecerem feedback em tempo real tanto aos alunos quanto aos professores. Isso não apenas auxilia os alunos a compreenderem seu progresso, mas também orienta os educadores sobre a eficácia de suas abordagens de ensino, possibilitando ajustes instantâneos para maximizar a eficiência do aprendizado (SCAICO; QUEIROZ; SCAICO, 2014).

Ao personalizar o aprendizado, o Big Data promove a autonomia do aluno em sua própria jornada educacional. Essa abordagem tende a aumentar a participação, visto que os materiais e atividades são adaptados aos interesses individuais, tornando o processo de aprendizagem mais relevante e envolvente.

3.4 Previsão de demandas futuras e alocação eficiente de recursos

O Big Data surge com uma ampla gama de ferramentas para analisar diversas tendências passadas e presentes, sendo crucial para as instituições educacionais na previsão de demandas por cursos, disciplinas ou programas. Essa capacidade preditiva, permite preparações antecipadas, assegurando a alocação de recursos de acordo com as necessidades futuras, evitando a subutilização ou superlotação dos recursos disponíveis. Ele também desempenha um papel crucial no planejamento orçamentário estratégico, permitindo que as instituições educacionais direcionem seus recursos financeiros de forma mais eficiente. Com uma compreensão clara das necessidades prioritárias, os investimentos podem ser mais focados nos aspectos que mais impactam a qualidade da educação (SCHROER, 2022).

Ademais, o estudo aprofundado de dados demográficos, socioeconômicos e geográficos dos alunos através do Big Data é fundamental para identificar desigualdades na distribuição de recursos. Essa análise promove uma alocação

mais justa e equitativa, direcionando recursos adicionais para regiões ou grupos com maior necessidade, fomentando um acesso mais igualitário à educação.

Além disso, a avaliação de padrões de desempenho do corpo docente e suas interações com os alunos, facilitada pelo Big Data, permite uma alocação mais adequada de professores, atendendo a necessidades específicas de disciplinas e grupos de estudantes. Essa estratégia maximiza a eficácia do ensino e garante que cada educador seja alocado onde sua expertise é mais necessária.

Em conclusão, além de fornecer um conjunto de ferramentas transformadoras no âmbito educacional, o Big Data possibilita o monitoramento constante e eficiente na alocação de recursos, permitindo ajustes rápidos e adaptações conforme as necessidades evoluem. Através da análise de tendências e padrões com tais ferramentas, as instituições podem prever necessidades futuras e endereçar desigualdades, garantindo que recursos sejam alocados onde são mais necessários, promovendo uma educação mais justa e acessível.

3.5 Aprimoramento contínuo do ensino

O Big Data viabiliza uma análise detalhada do desempenho dos alunos, identificando padrões, áreas de dificuldade e sucessos. Essa análise é fundamental para compreender onde e como os alunos estão aprendendo, fornecendo dados importantes para ajustar estratégias de ensino (SCAICO; QUEIROZ; SCAICO, 2014).

Os dados do Big Data são fundamentais também no desenvolvimento profissional dos educadores, auxiliando na identificação de áreas que poderiam melhorar e oferecendo caminhos para o aprimoramento de habilidades e práticas de ensino. A análise de dados educacionais provenientes de cursos online e outras fontes proporciona aos professores informações sobre o desempenho e padrões de aprendizagem dos alunos. Esses dados são valiosos para uma análise mais profunda das trajetórias de aprendizagem dos alunos e das metodologias de ensino. Técnicas de visualização de dados são particularmente úteis para ajudar os professores, especialmente aqueles com conhecimento disciplinar limitado, a interpretar e compreender melhor os dados dos alunos (LI; JIANG, 2021)

Em resumo, a aplicação do Big Data no aprimoramento contínuo do ensino oferece uma oportunidade para personalização, adaptação e melhoria das práticas de ensino, resultando em uma educação mais eficaz e adaptada às necessidades individuais dos alunos.

4 EXPERIÊNCIAS BEM-SUCEDIDAS

Nas seções seguintes iremos explorar e analisar diversos casos que ilustram as aplicações do Big Data no contexto educacional. Ao investigar experiências e resultados concretos, busca-se não apenas compreender a extensão do impacto dessas tecnologias, mas também destacar as potenciais transformações que podem moldar o futuro da educação.

4.1 Duolingo

O Duolingo se destaca como uma plataforma educacional de aprendizado de idiomas, notável por sua abordagem inovadora e eficaz. Fundada em 2011, a plataforma oferece uma diversidade de cursos em vários idiomas, empregando métodos interativos e lúdicos para engajar os usuários no processo de aprendizagem. A integração de tecnologia e dados é crucial no sucesso do Duolingo, com foco particular na personalização do ensino.

No cerne do Duolingo está a estratégica utilização de Big Data para customizar a experiência de aprendizado de cada usuário. A tecnologia de aprendizado de máquina própria do Duolingo, conhecida como Birdbrain, utiliza esses dados para adaptar o ensino à proficiência e às necessidades individuais de cada aluno, gerando lições dinâmicas e personalizadas. (BICKNELL; BRUST, 2020)

Essa tecnologia representa um exemplo significativo de como o Big Data está transformando a educação, especialmente no campo do aprendizado de línguas. Este modelo de aprendizado de máquina foi desenvolvido para aprender tanto dos usuários quanto sobre eles, visando aprimorar a experiência de aprendizado.

A essência do Birdbrain reside em sua capacidade de avaliar continuamente o quanto os aprendizes sabem, bem como o nível de dificuldade dos diferentes tipos de material linguístico. Ele faz uma previsão sobre a probabilidade de um aluno acertar um determinado exercício. Isso complementa o sistema de personalização existente, que determina o quanto um aprendiz conhece cada palavra específica, estendendo esse conhecimento para todas as facetas da linguagem na aplicação.

O Birdbrain é integrado ao "Session Generator" do Duolingo, que cria lições a partir de um vasto conjunto de exercícios potenciais usando um algoritmo sofisticado. Com o Birdbrain, o Session Generator pode determinar quais exercícios são do nível de dificuldade ideal para cada aluno específico, personalizando assim a experiência de aprendizado, permitindo que os alunos avancem rapidamente em áreas de proficiência e dediquem mais tempo aos tópicos mais desafiadores.

A personalização no Duolingo baseia-se na capacidade da plataforma de analisar esses dados de forma perspicaz. Utilizando algoritmos avançados, o Duolingo identifica as habilidades e dificuldades específicas de cada usuário, bem como o estilo de aprendizado mais eficaz para cada pessoa. Essa análise permite que a plataforma ajuste dinamicamente o conteúdo apresentado, adaptando o nível de dificuldade e oferecendo exercícios que correspondem ao ritmo e às preferências individuais dos usuários.

Além da personalização, o Duolingo utiliza o Big Data para avaliar continuamente a eficácia de seus métodos de ensino e ajustar suas estratégias pedagógicas. Com mais de 300 milhões de usuários completando mais de sete bilhões de exercícios de aprendizagem de idiomas por mês (DOL, 2021), a plataforma usa suas extensas bases de dados para informar melhorias na experiência do usuário e na infraestrutura subjacente, além de conduzir pesquisas regulares para otimizar práticas de ensino e resultados de aprendizagem, bem como para desenvolver novas funcionalidades. Isso é realizado por meio de uma abordagem de testes A/B, onde um conjunto de usuários da aplicação é designado ao grupo de controle (grupo A), enquanto outros são inseridos no grupo experimental (grupo B). O grupo A utiliza a versão atual da aplicação, e o grupo B experimenta a aplicação com as novas funcionalidades. Diariamente, o serviço de experimentos produz relatórios para cada experimento em andamento, conduzindo análises estatísticas em todas as métricas pertinentes àquela pesquisa a fim de criar gráficos para identificar o seu sucesso ou não. Ao se basear no resultado comparativo entre esses grupos, em quesitos como, por exemplo, lições completadas por parte dos usuários, é escolhida a versão que obteve mais resultados positivos. (APRAMEYA, 2020)

Um exemplo de funcionalidade que foi testada pela equipe do Duolingo foi para adicionar placares de liderança ao aplicativo. O resultado pode ser observado no gráfico 1.

Gráfico 1 - Resultado do teste realizado pelo Duolingo



Fonte: Duolingo Blog, 2020

A linha azul representa as pessoas que continuaram usando a versão atual do aplicativo, enquanto a verde representa as pessoas que passaram a usar a versão com a funcionalidade ativa.

Embora a inclusão de um placar de lideranças no aplicativo não seja uma melhoria diretamente no quesito educacional, pôde-se observar um aumento no número de aulas que os alunos iniciaram e também no número de aulas que concluíram. Essas métricas mostraram ganhos tanto em engajamento quanto em aprendizado, o que levou a equipe da plataforma a lançar o experimento para todos os usuários.

Existem diversas maneiras de despertar o interesse dos alunos e engajá-los nas atividades educacionais mesmo que as mudanças não sejam feitas diretamente no conteúdo ensinado. Essa técnica recebe o nome de gamificação e consiste na utilização de elementos como sistemas de recompensas, objetivos, níveis, narrativa, interação e muitos outros que normalmente são encontrados em jogos, com a intenção de engajar e motivar as pessoas que estão estudando da mesma maneira

que os jogos fazem (FARDO, 2013), transformando aquele momento de aprendizado em algo envolvente e divertido.

Em resumo, o Duolingo exemplifica como o Big Data e a gamificação alinhados são ferramentas poderosas que podem transformar a experiência de aprendizado, oferecendo personalização e eficácia. Ao analisar e compreender as nuances individuais dos usuários, a plataforma não só facilita o aprendizado de idiomas, mas também evolui continuamente para atender às necessidades dinâmicas de seus alunos.

4.2 Summit Public Schools

As Summit Public Schools, um grupo de escolas charter na área da Baía de São Francisco, Califórnia, Estados Unidos, utiliza intensamente o Big Data em suas práticas educacionais. O foco principal está no Plano de Aprendizado Personalizado, um painel para estudantes desenvolvido internamente, que fornece a professores e alunos acesso a uma ampla gama de dados sobre cada aluno. Este sistema permite o ajuste das instruções para atender às necessidades individuais dos alunos, como simplificar a linguagem em um teste de matemática para ajudar alunos que estão aprendendo inglês, mas entendem os conceitos matemáticos.

As escolas Summit desenvolvem suas próprias ferramentas, reformulam as existentes ou criam exibições de dados que incluem múltiplas fontes de informações, proporcionando aos alunos visões sobre o próprio desempenho. Um exemplo disso é a abordagem da Summit ao permitir que os alunos tenham acesso aos seus dados e ferramentas para trabalhar com eles. Os alunos fazem login diariamente em seu Plano de Aprendizado Personalizado para definir metas de aprendizado com a ajuda de seus professores, acompanhar o progresso, receber feedback e acessar recursos de aprendizado (GATES FOUNDATION, 2017).

O Plano de Aprendizado Personalizado (PAP) das Summit Public Schools representa um avanço significativo na educação personalizada, utilizando a tecnologia para melhorar o ensino e a aprendizagem. Este sistema oferece aos alunos a possibilidade de serem aprendizes autodirigidos, trabalhando com conteúdos em seu próprio ritmo e realizando avaliações conforme necessário. Além

disso, o PAP engaja os alunos em projetos de aprendizado profundo, permitindo que eles apliquem o conteúdo aprendido de maneira significativa e autêntica.

Para os professores, o PAP serve como uma ferramenta valiosa para personalizar o ensino, fornecendo dados em tempo real sobre o desempenho dos alunos. Isso permite que os educadores ajustem suas estratégias de ensino e forneçam suporte adicional, como mentoria e coaching. Além disso, a plataforma oferece um guia detalhado e recursos para ajudar os professores a implementar com sucesso o aprendizado personalizado em sala de aula. Os professores da Summit afirmam que os Planos de Aprendizado Personalizados transformaram a maneira como ensinam, facilitando a vida deles e dos alunos, tornando-os melhores educadores e os estudantes mais bem-sucedidos

Ao capacitar os alunos a assumirem o controle de seu próprio aprendizado e fornecer aos professores as ferramentas necessárias para apoiá-los, o PAP das Summit Public Schools destaca-se como um modelo exemplar de como a tecnologia pode ser utilizada para criar experiências educacionais mais eficazes e personalizadas.

A abordagem inovadora das Summit Public Schools em relação ao uso do Big Data na educação está intrinsecamente ligada ao seu modelo pedagógico. O foco nas práticas educacionais baseadas em dados, como evidenciado no Plano de Aprendizado Personalizado, complementa e fortalece os quatro pilares fundamentais do método de ensino da Summit: mentoria, projetos, autodireção e expedições (SUMMIT PUBLIC SCHOOLS, 2020?). Enquanto o Big Data possibilita um ensino mais personalizado e adaptado às necessidades individuais de cada aluno, os métodos de ensino da Summit, incluindo a aprendizagem baseada em projetos e o ciclo de aprendizado auto-dirigido, garantem que os alunos não apenas absorvam conhecimento, mas também desenvolvam habilidades cruciais para o sucesso no mundo real. Essa sinergia entre a utilização de dados e a abordagem pedagógica coloca as Summit Public Schools na vanguarda da inovação educacional, preparando os estudantes não só academicamente, mas também para os desafios e oportunidades da vida fora da escola.

Os quatro pilares do método de ensino das Summit Public Schools exemplificam a integração eficaz do Big Data na educação.

- Mentoria personalizada: Através do uso de dados sobre o progresso e as preferências de aprendizado de cada aluno, permitindo uma

orientação mais precisa e uma maior preparação para a faculdade. Um estudante do segundo ano de uma unidade escolar relata a importância da existência de mentores selecionados (HU *et al.*, 2020)

- **Projetos:** Utilizando-se do Big Data para personalizar o conteúdo educacional, adaptando-se ao ritmo e necessidade do estudante.
- **Autodireção:** Onde os estudantes têm a autonomia para escolher os recursos que preferem, decidir o momento ideal para realizar avaliações e identificar as condições ambientais que mais favorecem seu aprendizado. Esta abordagem valoriza a independência do aluno, permitindo-lhes moldar sua jornada educacional de acordo com suas preferências e ritmos individuais.
- **Expedições:** que se beneficiam do Big Data ao oferecer uma gama diversificada de cursos e experiências práticas, baseando-se em análises de interesses e tendências dos alunos para fornecer oportunidades de aprendizado relevantes e envolventes. Um estudante, por exemplo, criou um projeto de podcast com essa iniciativa (HU *et al.*, 2020)

Ao unir o Big Data com sua pedagogia inovadora, as Summit Public Schools exemplificam como a tecnologia de dados pode transformar a educação, proporcionando um ambiente de aprendizado mais personalizado, interativo e eficaz para os alunos.

4.3 Khan Academy

A Khan Academy, uma renomada organização educacional sem fins lucrativos, é reconhecida por sua plataforma online que oferece uma ampla variedade de recursos educacionais em diversas disciplinas. Sua proeminência no cenário educacional contemporâneo está atrelada à sua missão de fornecer educação de qualidade acessível a todos, independentemente de sua localização geográfica. Os recursos da Khan Academy, que incluem vídeos educativos, exercícios práticos e ferramentas de feedback, solidificam sua posição como uma referência em educação online.

A missão da Khan Academy de oferecer educação de excelência que transcenda barreiras geográficas e econômicas é evidente em sua abordagem pedagógica. A plataforma adota um modelo de aprendizado personalizado, permitindo que os alunos progridam de acordo com seus ritmos individuais. A capacidade de adaptar exercícios com base no desempenho de cada aluno é um aspecto chave dessa abordagem, fornecendo suporte adicional nas áreas necessárias.

A singularidade da Khan Academy também se manifesta no uso estratégico de Big Data para aprimorar continuamente a plataforma e otimizar a experiência de aprendizado individual. A organização emprega inteligência artificial e análise de dados para entender os padrões de aprendizagem de cada aluno, oferecendo conteúdos e correções personalizados. Esse uso de dados para melhorias constantes é fundamental para a missão da Khan Academy de promover uma educação personalizada e eficaz.

Além disso, a Khan Academy utiliza Big Data para melhorar a excelência das atividades e a experiência de aprendizado. A organização emprega análises comportamentais para entender a qualidade das aulas, monitorando o engajamento dos alunos com os vídeos e os resultados dos testes. A Khan Academy também usa Big Data para testar elementos da interface do usuário (UI/UX), um fator crítico para o engajamento e retenção dos alunos. A eficácia de cada design é determinada com base no tempo gasto no site, nas ações realizadas nas lições, nos cliques e nas taxas de rejeição.

Adicionalmente, a utilização de Big Data pela Khan Academy permite uma abordagem mais fundamentada e orientada a dados na identificação das necessidades educacionais dos alunos. Ao analisar o desempenho dos alunos em tempo real, a Khan Academy pode identificar áreas de dificuldade, permitindo uma intervenção mais rápida e eficaz. Isso não apenas melhora a eficiência do processo de aprendizado, mas também ajuda a manter os alunos engajados e motivados.

A Khan Academy também se destaca pelo seu compromisso com a inovação contínua. A organização não se limita ao uso de tecnologias existentes, mas busca ativamente desenvolver novas abordagens e soluções que possam melhorar ainda mais a experiência educacional. Isso inclui a experimentação com novos formatos de conteúdo, técnicas de ensino e estratégias de avaliação, sempre com o objetivo de tornar a educação mais acessível, inclusiva e eficaz.

Em um cenário educacional global cada vez mais digitalizado e centrado em dados, a Khan Academy serve como um modelo exemplar de como o Big Data pode ser aplicado para enriquecer e personalizar a experiência de aprendizagem. Suas práticas inovadoras e sua abordagem orientada a dados oferecem insights valiosos para educadores, desenvolvedores de tecnologia educacional e formuladores de políticas, demonstrando o potencial transformador do Big Data na educação.

A plataforma é uma das principais fontes de aprendizado online confiáveis por professores e alunos nos EUA, consistentemente gerando resultados que superam as expectativas de crescimento típicas. Um estudo conduzido pela Khan Academy procurou responder a uma pergunta crítica na tecnologia educacional de como a tecnologia realmente poderia ajudar no aprendizado dos alunos e a abordagem utilizada foi a comparação dos alunos com seu próprio desempenho passado ao longo de dois anos letivos, monitorando como as mudanças no uso da Khan Academy afetavam seu desempenho. Os resultados do estudo foram positivos, mostrando que os alunos que aumentaram o uso da Khan Academy de um ano para o outro experimentaram melhorias significativas em seus escores no teste MAP Growth, um teste realizado para medir o progresso do aluno nas escolas K-12 (expressão usada para designar intervalo em anos que envolve o ensino primário e secundário nas escolas americanas). O estudo utilizou um Índice de Crescimento Condicional (CGI) para rastrear a evolução dos alunos, que leva em conta fatores como pontuações iniciais e níveis de série, tornando-o um indicador confiável de progresso real de aprendizado. Um exemplo do estudo mostra que até mesmo um pequeno aumento no uso da Khan Academy poderia resultar em um aumento significativo nos escores CGI. O design do estudo controlou variáveis como motivação ou status socioeconômico, já que cada aluno era essencialmente seu próprio grupo de controle (NWEA, 2023).

Este estudo destaca o potencial da Khan Academy como uma ferramenta valiosa para aprimorar os resultados de aprendizado, mesmo em contextos desafiadores como, por exemplo, a pandemia do Covid-19. Ele sublinha a eficácia da plataforma em impulsionar o desempenho acadêmico e sugere que qualquer aumento no uso, mesmo que pequeno, provavelmente beneficiará o aprendiz.

Em suma, a Khan Academy não só cumpre sua missão de fornecer educação acessível e de qualidade, mas também demonstra um compromisso com a inovação educacional por meio da aplicação de Big Data. A organização exemplifica como a

análise de dados pode atuar como um catalisador para o desenvolvimento constante e a personalização eficaz do processo de aprendizado, estabelecendo-se como uma líder no avanço da educação online e na promoção de experiências educacionais mais eficientes e individualizadas.

4.4 Purdue University

Outro caso de estudo a ser citado no que se refere à inserção do big data e suas ferramentas na educação é a Purdue University, nos Estados Unidos, que desenvolveu o sistema “Course signals” que, através do uso de modelos preditivos, coleta e analisa bases de dados educacionais com o intuito de determinar quais estudantes se encontram em risco de reprovação em alguma matéria, por exemplo. Levando em conta quatro fatores: performance, esforço, histórico acadêmico e características dos estudantes, o algoritmo do sistema Course signal calcula e fornece resultados baseados na representação de um sinal de trânsito. Se um aluno está em condição de alto risco de reprovação, um sinal vermelho é mostrado na página principal do aluno, o amarelo significa um potencial risco para dar atenção e finalmente, o sinal verde indica uma grande chance de aprovação com sucesso (ARNOLD; PISTILLI, 2012).

Com base nesse sistema, professores podem criar programas de intervenção, calendários de estudo personalizados para cada aluno, indicar materiais específicos entre outras soluções.

Uma pesquisa foi conduzida para analisar a taxa de retenção de alunos que ingressaram na Purdue University através dos anos e que se matricularam e frequentavam cursos que possuíam a presença do sistema do Course Signals.

Tabela 1 - Taxa de retenção do grupo ingressante na universidade Purdue em 2007

Número de cursos que utilizam o Course signals	Tamanho do grupo	Ano 1	Ano 2	Ano 3	Ano 4
Sem “Course Signals”	5,134	83.44%	73.14%	70.47%	69.40%

No mínimo 1 curso	1,518	96.71%	94.73%	90.65%	87.42%
1 curso	1,311	96.57%	94.13%	89.70%	86.50%
2 ou mais cursos	207	97.58%	98.55%	96.62%	93.24%

Fonte: Arnold & Pistilli, 2012

Tabela 2 - Taxa de retenção do grupo ingressante na universidade Purdue em 2008

Número de cursos que utilizam o Course signals	Tamanho do grupo	Ano 1	Ano 2	Ano 3
Sem Course Signals	4,221	81.69%	75.08%	73.21%
No mínimo 1 curso	2,690	96.25%	89.55%	85.17%
1 curso	2,125	95.62%	88.00%	83.58%
2 ou mais cursos	565	98.58%	95.40%	91.15%

Fonte: Arnold & Pistilli, 2012

Tabela 1 - Taxa de retenção do grupo ingressante na universidade Purdue em 2009

Número de cursos que utilizam o Course signals	Tamanho do grupo	Ano 1	Ano 2
Sem Course Signals	3,164	87.67%	81.89%
No mínimo 1 curso	2,962	90.34%	83.22%
1 curso	2,296	87.72%	80.87%
2 ou mais cursos	666	99.40%	91.44%

Fonte: Arnold & Pistilli, 2012

Os resultados da pesquisa, verificados nas tabelas 1, 2 e 3 indicam que houve uma taxa de retenção maior por parte dos alunos ingressantes, em cursos que se utilizavam do sistema do Course Signals. Tais dados comprovam a eficácia da utilização do big data e suas ferramentas para melhorar a educação na Purdue University.

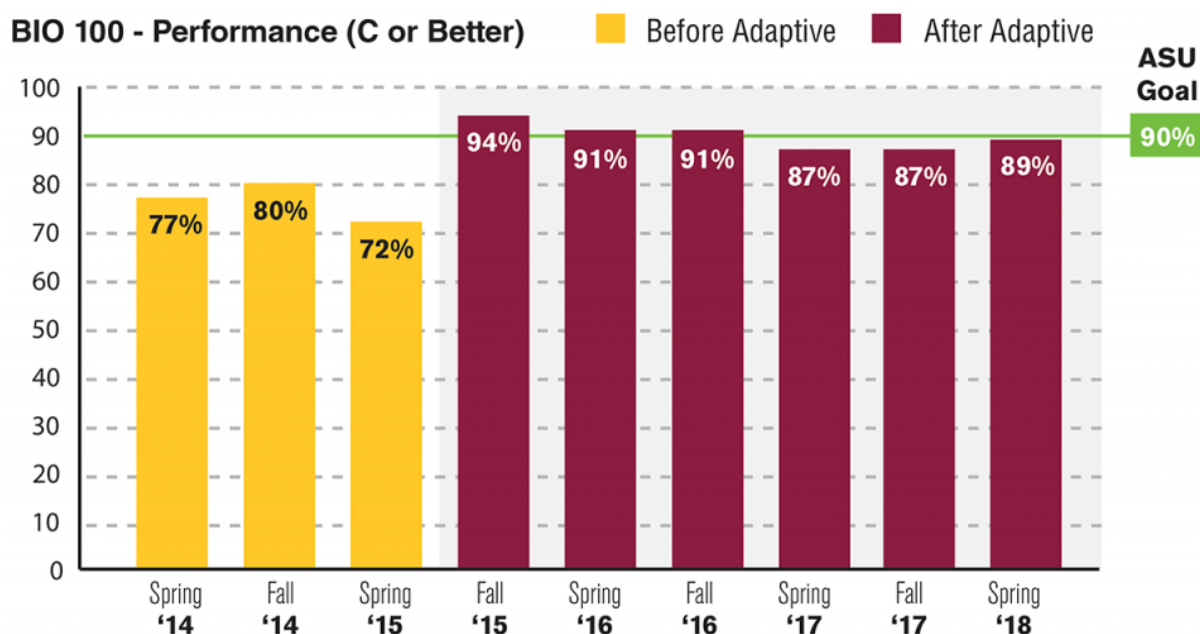
4.5 Arizona State University

A Arizona State University é outro exemplo de instituição que está investindo na personalização do aprendizado e buscando inovar suas salas de aula. “Costumávamos ensinar a mesma coisa a todos ao mesmo tempo. Agora, estamos conectando o aluno certo à lição certa. Estamos mudando a estrutura do ensino superior de estática para dinâmica.” diz Dale Johnson, diretor de iniciativas de aprendizagem adaptativa da EdPlus, uma unidade empresarial da Arizona State University responsável por modelos digitais de ensino e aprendizagem para reduzir os obstáculos educacionais e aumentar o sucesso dos alunos no ensino superior.

Como teste inicial, alunos não graduados matriculados em cursos de biologia para cumprir um requisito de ciências em suas especialidades foram os primeiros a experimentar o formato de aprendizagem adaptativa. Esse segmento é conhecido por seu baixo desempenho e altas taxas de evasão. Basicamente, o livro didático foi substituído por uma experiência online semelhante a um livro didático, mas que atende às necessidades do aluno. Os resultados após um semestre surpreenderam bastante. A taxa de aprovação teve um aumento de 24% e a taxa de evasão diminuiu em incríveis 90%. “Este nível de sucesso resultante de um investimento tão modesto convenceu-nos de que estávamos a fazer algo certo. Esses resultados foram replicados em outras escolas, ao longo de vários anos, e foram validados por estudos independentes em larga escala.” disse Jim Thompson, CEO da CogBooks, empresa parceira na criação da tecnologia.

Após o sucesso do primeiro teste, os responsáveis pelo projeto continuaram a aprimorar o programa. Em seguida, foi a vez de Carlo Maley, professor associado da escola, se voluntariar para testar o programa com seus alunos. O livro didático foi modificado e sua versão traduzida foi integrada à plataforma de aprendizagem adaptativa CogBooks. Posteriormente, houve uma inversão no formato da sala de aula, com a entrega de instruções online fora do ambiente presencial, enquanto as atividades interativas e práticas eram conduzidas durante o tempo de aula.

Gráfico 2 - Resultado do segundo teste da plataforma



Fonte: ASU News, 2019

As barras representam a performance dos alunos por semestre. As amarelas antes de utilizar a plataforma de aprendizagem adaptativa e as roxas após a implementação do programa. Os resultados mostraram uma melhora significativa na performance dos alunos. “Ao olharmos para os nossos sucessos até agora, já vemos maior persistência e maior desempenho dos nossos alunos. Este programa dinâmico é o futuro da educação, e a ASU (Arizona State University) é líder global neste esforço”, disse Dale Johnson.

E o programa não se limita a esses dados, os responsáveis continuam buscando aprimorar cada vez mais a tecnologia e promover uma evolução constante no método de ensino da instituição (LEANDER, 2019).

5 DESAFIOS DO BIG DATA

No cenário atual, a implementação e gestão eficazes do Big Data estão intrinsecamente associadas a uma série de desafios complexos e multifacetados.

O volume massivo de dados gerados continuamente representa uma das principais barreiras a serem superadas. A necessidade de infraestruturas escaláveis, capazes de armazenar e processar essa grande quantidade de informações, é fundamental para o funcionamento dos projetos de Big Data (GORTON; KLEIN, 2015). O custo associado ao desenvolvimento e manutenção dessas infraestruturas constitui uma barreira financeira considerável. O investimento necessário em armazenamento, processamento e tecnologias analíticas pode ser substancial.

Questões relacionadas a velocidade de processamento, diversidade, qualidade e veracidade dos dados também são importantes pois exigem sistemas ágeis e eficientes (KOLAJO; DARAMOLA; ADEBIYI, 2019).

A integração de sistemas e tecnologias também é um desafio significativo, especialmente considerando a diversidade de plataformas e ferramentas disponíveis. A capacidade de extrair informações úteis e interpretar padrões complexos requer habilidades analíticas avançadas, nos levando a outro problema, a escassez de especialistas, que representa um desafio persistente, uma vez que a demanda por profissionais qualificados supera a oferta (OLIVEIRA, 2020).

As implicações éticas relacionadas à utilização de Big Data estão cada vez mais evidentes, com preocupações centradas na privacidade, transparência e segurança. A proteção de informações sensíveis e a adesão às normas de privacidade demandam a implementação de estratégias de segurança e conformidade. Especialmente na coleta e análise de dados massivos no quesito educacional, que frequentemente envolvem informações pessoais sensíveis de estudantes, assegurar a privacidade desses dados é imperativo para prevenir violações éticas. Nesse contexto, as instituições têm a responsabilidade de garantir que os resultados provenientes das análises sejam utilizados de maneira íntegra e ética.

A Lei Geral de Proteção de Dados (BRASIL, 2018), também conhecida como LGPD, (Lei nº 13.709/2018) desempenha um papel significativo na abordagem dessas questões éticas. Ao estabelecer diretrizes específicas para o tratamento de

dados pessoais, a LGPD contribui para a mitigação de preocupações relacionadas à privacidade. Ela impõe a necessidade de consentimento informado e esclarecido, exigindo que as instituições obtenham a aprovação explícita dos alunos ou de seus responsáveis antes de realizar a coleta e processamento de dados. Além disso, a LGPD reforça a importância da transparência ao impor a divulgação clara sobre a finalidade e escopo da coleta de dados. As instituições são obrigadas a adotar práticas que garantam a segurança e a conformidade com padrões éticos na manipulação de informações sensíveis. Abordar essas questões requer uma combinação de avanços tecnológicos, políticas regulatórias eficazes e considerações éticas aprimoradas. Essa compreensão profunda dos desafios do Big Data é essencial para orientar melhores práticas no contexto da gestão e aplicação dessa significativa quantidade de dados.

6 CONSIDERAÇÕES FINAIS

Este trabalho buscou analisar a influência do big data na área da educação, se utilizando de ferramentas desenvolvidas por instituições com o intuito de transformar esse grande volume de dados em informação de valor para fins de aprendizagem. Além disso, procurou-se também considerar o big data e as ferramentas de análises de dados como uma opção viável para substituir metodologias de ensino que podem ser consideradas ultrapassadas. Para isso foram analisados estudos de referências com base em conceitos relacionados ao tema.

No que tange à questão do paralelo traçado pelo big data e suas aplicações na educação, pode-se concluir que a partir do uso da análise de dados, pontos positivos fundamentais relacionados à evolução do processo de aprendizagem puderam ser notados, como a melhora na personalização de aprendizado de um estudante, levando em conta não apenas seu desempenho acadêmico como também seus interesses, e a na evasão de alunos, com base em um volume de dados embasado em condições socioeconômicas e emocionais.

Este estudo destacou a influência significativa do big data na transformação da educação através da observação de plataformas como Duolingo e Khan Academy, bem como instituições como as Summit Public Schools, Purdue University e Arizona State University. Após análise, ficou evidente que o big data pode ser um grande aliado, auxiliando e aprimorando diversos setores na educação.

Em contrapartida, ao longo da análise do tema, foram detalhadas as dificuldades que podem se tornar empecilhos para o desenvolvimento do big data nessa área. Mesmo que existam ferramentas que lidem com o grande volume de dados, eles se encontram em contínuo crescimento e o desenvolvimento de uma infraestrutura se torna necessária para que se acompanhe tal fenômeno, questões como velocidade de processamento de dados e ainda privacidade dos dados são exemplos de pontos negativos que foram encontrados no estudo.

Por fim, apesar dos desafios relacionados à infraestrutura e privacidade de dados, o potencial do big data para promover uma solução para a questão de metodologias de ensino ultrapassadas, revitalizar métodos educacionais e possibilitar um ensino mais adaptativo e eficaz é inegável. Portanto, a integração do

big data na educação representa um passo promissor para um futuro no qual a aprendizagem é mais adaptada às necessidades individuais dos alunos, contribuindo para um sistema educacional mais eficiente e inclusivo.

REFERÊNCIAS

- ALVES, C.; LAVOR, L.; PEREIRA, H. **Evasão escolar: um desafio para a educação na atualidade** | *Revista de Pesquisa Interdisciplinar*. 2017. Disponível em: <<https://doi.org/10.24219/rpi.v2i1.132>> Acesso em: 10 dez. 2023.
- APRAMEYA, L. **Improving Duolingo, one experiment at a time**. 2020. Duolingo Blog. Disponível em: <<https://blog.duolingo.com/improving-duolingo-one-experiment-at-a-time/>> Acesso em: 10 dez. 2023.
- ARNOLD, K.; PISTILLI, M. **Course signals at Purdue: Using learning analytics to increase student success**. 2012. ACM International Conference Proceeding Series. Disponível em: <https://www.researchgate.net/publication/254462830_Course_signals_at_Purdue_Using_learning_analytics_to_increase_student_success> Acesso em: 10 dez. 2023.
- BERISHA, B.; MĚZIU, E.; SHABANI, I. **Big data analytics in Cloud computing: an overview - Journal of Cloud Computing**. 2022. SpringerOpen. Disponível em: <<https://doi.org/10.1186/s13677-022-00301-w>> Acesso em: 10 dez. 2023.
- BICKNELL, K.; BRUST, C. **Learning how to help you learn: Introducing Birdbrain!** 2020. Duolingo Blog. Disponível em: <<https://blog.duolingo.com/learning-how-to-help-you-learn-introducing-birdbrain/>> Acesso em: 10 dez. 2023.
- BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais**. Diário Oficial da União, Brasília, DF, 15 ago. 2018. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>. Acesso em: 10 dez. 2023.
- COULOURIS, G. *et al.* **Distributed systems: concepts and design**. 2005. Disponível em: <<http://bedford-computing.co.uk/learning/wp-content/uploads/2016/03/george-coulouris-distributed-systems-concepts-and-design-5th-edition.pdf>> Acesso em: 10 dez. 2023.
- DOL, Q. **How Duolingo Builds Its Data Science Methodology**. 2021. Built In. Disponível em: <<https://builtin.com/data-science/how-duolingo-builds-data-science-methodology>> Acesso em: 10 dez. 2023.

DUCANGE, P.; FAZZOLARI, M.; MARCELLONI, F. **An overview of recent distributed algorithms for learning fuzzy models in Big Data classification - Journal of Big Data**. 2020. SpringerOpen. Disponível em: <<https://doi.org/10.1186/s40537-020-00298-6>> Acesso em: 10 dez. 2023.

ESTEVAM, P. **Aprendizagem personalizada: definições, benefícios e 3 dicas de como trabalhá-la**. 2021. Rubeus. Disponível em: <<https://rubeus.com.br/blog/aprendizagem-personalizada/>>. Acesso em: 10 dez. 2023.

FARDO, M. **A gamificação aplicada em ambientes de aprendizagem**. 2013. Revista Novas Tecnologias na Educação, Porto Alegre, v. 11, n. 1. Disponível em: <<https://seer.ufrgs.br/index.php/renote/article/view/41629>>. Acesso em: 10 dez. 2023.

FERREIRA, J. **Big data in education: The five types that matter**. 2013. Disponível em: <<https://publicservicesalliance.org/wp-content/uploads/2013/07/Big-Data-in-Education-The-5-Types-That-Matter-Knewton-Blog.pdf>> Acesso em: 10 dez. 2023.

GATES FOUNDATION. **Reaching the Summit of Data-Driven Instruction**. 2017. Disponível em: <<https://usprogram.gatesfoundation.org/news-and-insights/articles/reaching-the-summit-of-data-driven-instruction>> Acesso em: 10 dez. 2023.

GORTON, I.; KLEIN, J. **Distribution, Data, Deployment: Software Architecture Convergence in Big Data Systems**. 2015. Disponível em: <<https://doi.org/10.1109/MS.2014.51>> Acesso em: 10 dez. 2023.

HAN, J.; KAMBER, M.; PEI, J. **Data Mining: Concepts and Techniques**. 2011. Disponível em: <<https://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>> Acesso em: 10 dez. 2023.

HARIRI, R.; FREDERICKS, E.; BOWERS, K. **Incerteza na análise de big data: pesquisa, oportunidades e desafios**. 2019. Disponível em: <<https://doi.org/10.1186/s40537-019-0206-3>> Acesso em: 10 dez. 2023.

HU, E. *et al.* **Students reflect on how effective the Summit Model is in their education**. 2020. Summit News. Disponível em:

<<https://summitpsnews.org/2020/03/27/students-reflect-on-how-effective-the-summit-model-is-in-their-education/>> Acesso em: 10 dez. 2023.

KOLAJO, T.; DARAMOLA, O.; ADEBIYI, A. **Big data stream analysis: a systematic literature review**. 2019. Disponível em: <<https://doi.org/10.1186/s40537-019-0210-7>> Acesso em: 10 dez. 2023.

LEANDER, S. **ASU develops world's first adaptive-learning biology degree**. 2019. ASU News. Disponível em: <<https://news.asu.edu/20190820-solutions-asu-develops-world-first-adaptive-learning-biology-degree>> Acesso em: 10 dez. 2023.

LEE, S.; CHUNG, J. **The Machine Learning-Based Dropout Early Warning System for Improving the Performance of Dropout Prediction**. 2019. Appl. Sci. 9, 3093. Disponível em: <<https://doi.org/10.3390/app9153093>> Acesso em: 10 dez. 2023.

LI, J.; JIANG, Y. **The Research Trend of Big Data in Education and the Impact of Teacher Psychology on Educational Development During COVID-19: A Systematic Review and Future Perspective**. 2021. Disponível em: <<https://doi.org/10.3389/fpsyg.2021.753388>> Acesso em: 10 dez. 2023.

MAHESH, B. **Machine Learning Algorithms - A Review**. 2019. Disponível em: <https://www.researchgate.net/publication/344717762_Machine_Learning_Algorithms_-_A_Review> Acesso em: 10 dez. 2023.

NWEA. **Elevating Math Scores: The Ongoing Success of MAP Accelerator**. 2023. Disponível em: <<https://www.nwea.org/resource-center/resource/elevating-math-scores-the-ongoing-success-of-map-accelerator/>> Acesso em: 10 dez. 2023.

OLIVEIRA, R. **A escassez de talentos de inteligência e dados**. 2020. ABRACD - ASSOCIAÇÃO BRASILEIRA DE CIÊNCIA DE DADOS. Disponível em: <<https://abracd.org/a-escassez-de-talentos-de-inteligencia-e-dados/>> Acesso em: 10 dez. 2023.

SABHARWAL, R., MIAH, S. **A new theoretical understanding of big data analytics capabilities in organizations: a thematic analysis**. 2021. J Big Data 8, 159. Disponível em: <<https://doi.org/10.1186/s40537-021-00543-6>> Acesso em: 10 dez. 2023.

SAGIROGLU, S; SINANC, D. **Big data: A review**. 2013. International Conference on Collaboration Technologies and Systems (CTS), San Diego, CA, USA. Disponível em: <<https://doi.org/10.1109/CTS.2013.6567202>> Acesso em: 10 dez. 2023.

SANTOS, V. **Estudo sobre um modelo de educação ultrapassado**. 2016. Anais III CONEDU. Campina Grande: Realize Editora. Disponível em: <<https://editorarealize.com.br/artigo/visualizar/21011>>. Acesso em: 10 dez. 2023.

SCAICO, P.; QUEIROZ, R.; SCAICO, A. **O conceito big data na Educação**. 2014. Disponível em: <https://www.researchgate.net/publication/300116871_O_conceito_big_data_na_Educacao> Acesso em: 10 dez. 2023.

SCHROER, A. **Big Data in Education: 10 Companies Delivering Insights to the Classroom**. 2022. Built In. Disponível em: <<https://builtin.com/big-data/big-data-in-education>> Acesso em: 10 dez. 2023.

SUMMIT PUBLIC SCHOOLS. **The Summit Model - Summit Public Schools**. [2020?]. Disponível em: <<https://summitps.org/the-summit-model/>> Acesso em: 10 dez. 2023.

UNICEF. **Educação brasileira em 2022 – a voz de adolescentes**. 2022. Disponível em : <<https://www.unicef.org/brazil/relatorios/educacao-brasileira-em-2022-a-voz-de-adolescentes>> Acesso em: 10 dez. 2023.

WAGNER, E.; ICE, P. **Data Changes Everything: Delivering on the Promise of Learning Analytics in Higher Education**. 2012. EDUCAUSE Review. Disponível em: <<https://er.educause.edu/articles/2012/7/data-changes-everything-delivering-on-the-promise-of-learning-analytics-in-higher-education>> Acesso em: 10 dez. 2023.

WAHEED, H. *et al.* **Predicting academic performance of students from VLE big data using deep learning models**. 2019. ScienceDirect. Disponível em: <<https://doi.org/10.1016/j.chb.2019.106189>> Acesso em: 10 dez. 2023.

YAĞCI, M. **Educational data mining: prediction of students' academic performance using machine learning algorithms - Smart Learning Environments**. 2022. SpringerOpen. Disponível em: <<https://doi.org/10.1186/s40561-022-00192-z>> Acesso em: 10 dez. 2023.

YOUNAS, M. **Research challenges of big data**. 2019. Disponível em:
<<https://doi.org/10.1007/s11761-019-00265-x>> Acesso em: 10 dez. 2023.