

BIG DATA E MACHINE LEARNING EM HOSPITAIS: avanços significativos na gestão hospitalar.***BIG DATA AND MACHINE LEARNING IN HOSPITALS: significant advances in hospital management.***

Marcelo Carletto Junior – marcelo.carletto@fatec.sp.gov.br
Faculdade de Tecnologia de Taquaritinga (Fatec) – Taquaritinga – SP – Brasil

Diego Aparecido Guariz – diego.guariz@fatec.sp.gov.br
Faculdade de Tecnologia de Taquaritinga (Fatec) – Taquaritinga – SP – Brasil

DOI: 10.31510/infa.v22i2.2338

Data de submissão: 25/09/2025

Data do aceite: 02/12/2025

Data da publicação: 20/12/2025

RESUMO

Nos últimos anos, o *Big Data* tem se tornado cada vez mais dominante na área da saúde, devido a três razões principais: a enorme quantidade de dados disponíveis, o aumento dos custos com saúde e a atenção personalizada. O processamento de *Big Data* na área da saúde se refere à geração, coleta, análise e armazenamento de dados clínicos muito vastos ou complexos para serem inferidos por métodos clássicos de processamento de dados. As fontes de *Big Data* para a área da saúde incluem a Internet das Coisas (IoT), o Prontuário Eletrônico do Paciente (PEP), que contém o histórico médico do paciente, diagnósticos, medicamentos, planos de tratamento, alergias, resultados de exames laboratoriais e de exames, sequenciamento genômico, imagens médicas, planos de saúde e outros dados clínicos. A metodologia deste estudo segue uma abordagem estruturada para revisar o impacto de *Big Data* e *Machine Learning* (ML) na gestão de informações em saúde.

Palavras-chave: Área da Saúde. *Big Data*. Prontuário Eletrônico. *Machine Learning*. Internet das Coisas. Gestão da Informação em Saúde.

ABSTRACT

In recent years, *Big Data* has become increasingly dominant in healthcare, driven by three main factors: the sheer volume of available data, rising healthcare costs, and the demand for personalized care. *Big Data* processing in healthcare refers to the generation, collection, analysis, and storage of clinical data too vast or complex to be processed by traditional data processing methods. Sources of *Big Data* for healthcare include the Internet of Things (IoT), Electronic Health Records (EHRs), which contain patient medical history, diagnoses, medications, treatment plans, allergies, laboratory and test results, genomic sequencing, medical images, health insurance plans, and other clinical data. This study's methodology

follows a structured approach to review the impact of *Big Data* and *Machine Learning* (ML) on healthcare information management.

Keywords: Healthcare. *Big Data*. Electronic Medical Records. *Machine Learning*. Internet of Things. Health Information Management.

1 INTRODUÇÃO

Nos últimos anos, o volume e a variedade de dados gerados transformaram drasticamente inúmeros campos, com impacto particular nos setores biomédico e de saúde. O *Big Data*, impulsionado pelo fluxo de informações de fontes como a internet, plataformas em nuvem e biossensores, está revolucionando o cenário moderno da informação (HUANG e DESTTECH, 2018). Iniciativas globais estão alavancando estrategicamente o *Big Data* para aprimorar insights e eficiência, com empresas como o Wal-Mart e o Facebook processando regularmente petabytes de dados (SAGIROGLU e SINANC, 2013). No entanto, o gerenciamento desses enormes conjuntos de dados vai além do armazenamento, abrangendo a extração de insights significativos e aplicáveis. Tecnologias como o Apache Spark, desenvolvido em Hadoop, surgiram para atender a essas crescentes demandas de dados (TAHA, 2025). Além disso, mecanismos de busca adaptados para *Big Data*, incluindo plataformas especializadas de busca biomédica, estão surgindo para ajudar médicos e pesquisadores a localizar rapidamente informações relevantes em vastos conjuntos de dados, aumentando ainda mais a eficiência e a precisão da análise de dados de saúde (MAHADEVKAR et al, 2022). O setor de saúde, particularmente nos EUA, enfrenta o desafio de gerenciar extensos conjuntos de dados médicos enquanto luta com barreiras econômicas e de infraestrutura.

À medida que a inteligência artificial (IA) amadurece por meio do progresso tecnológico contínuo e capacidades de processamento aprimoradas, ela transita de um conceito teórico para uma ferramenta prática de alto impacto em todos os domínios. Um ramo proeminente da IA, o

Machine Learning (ML), também conhecido como Aprendizado de Máquina (ML), concentra-se na automatização da análise de dados para descobrir padrões em grandes conjuntos de dados.

Essa capacidade tem se mostrado inestimável em áreas como análise de dados do consumidor e, cada vez mais, em aplicações biomédicas e de saúde, onde o poder preditivo do ML pode fornecer insights acionáveis. Eric Topol, uma figura de destaque na saúde digital, enfatizou que o futuro da medicina verá a IA em sinergia com o intelecto humano para elevar a prestação de serviços de saúde. A implantação do ML em análises de *Big Data* para a saúde tem o potencial de transformar dados em resultados significativos para a saúde, orientando médicos, aprimorando o atendimento ao paciente e otimizando os serviços de saúde (TAHA, 2025).

Este artigo tem como objetivo examinar a aplicação do *Machine Learning* (ML) no contexto do *Big Data* na área da saúde. A discussão apresentada contempla análises e avaliações que evidenciam tanto as potencialidades quanto as limitações dos diferentes métodos do *Machine Learning* (ML), com a finalidade de apoiar pesquisadores e profissionais em seus processos de tomada de decisão.

A metodologia deste estudo seguiu uma abordagem estruturada para revisar o impacto de *Big Data* e *Machine Learning* (ML) na gestão de informações em saúde. Foi realizada uma revisão sistemática da literatura, com busca em bases de dados eletrônicas utilizando palavras-chave relacionadas a "*big data*", "*machine learning*" e "gestão de informações em saúde". Os critérios de inclusão abrangeram publicações revisadas por pares a partir de 2010, com foco na aplicação prática dessas tecnologias em hospitais, clínicas e instituições de pesquisa.

Como justificativa, ressalta-se que os avanços recentes nessas técnicas podem ampliar significativamente o uso do *Big Data* na área da saúde, além de apontar caminhos promissores para futuras investigações científicas no campo.

2 BIG DATA E MACHINE LEARNING (ML) NA SAÚDE

A saúde entrou em uma era transformadora, marcada pelo uso crescente de tecnologias digitais para aprimorar o atendimento médico e a eficiência operacional. Uma das áreas mais proeminentes dessa transformação é a adoção de *Big Data* e *Machine Learning* (ML), que se tornaram ferramentas indispensáveis para os profissionais de saúde (Belle et al., 2015). Essas tecnologias permitem a análise de grandes conjuntos de dados, aprimorando a capacidade de descobrir padrões e gerar modelos preditivos que subsidiam uma melhor tomada de decisão em ambientes clínicos e administrativos. Por exemplo, Bates et al. (2014) destacam o potencial do

Big Data na saúde, observando que as vastas quantidades de dados geradas a partir de prontuários eletrônicos de saúde (PEs), estudos genômicos e dispositivos vestíveis podem ser processadas para fornecer insights que aprimoram os resultados dos pacientes. À medida que as ferramentas digitais se tornam mais integradas à saúde, o papel do *Big Data* e do *Machine Learning* na condução desses avanços continua a se expandir.

O uso de *Big Data* e *Machine Learning* (ML) tem se consolidado como essencial para a evolução da saúde, oferecendo amplo potencial para melhorar tanto os resultados clínicos quanto os administrativos. O *Big Data* em saúde envolve grandes volumes de informações provenientes de prontuários eletrônicos, dispositivos vestíveis e exames de imagem, cujo desafio está na gestão e análise eficiente para geração de insights (Ishwarappa & Anuradha, 2015). O *Machine Learning* (ML), por sua vez, possibilita que computadores aprendam e façam previsões a partir desses dados, apoiando decisões clínicas baseadas em evidências (Chawla & Davis, 2013).

Avanços em poder computacional e algoritmos de IA têm impulsionado a adoção dessas tecnologias, aplicadas em áreas como medicina personalizada, análise preditiva e sistemas de suporte à decisão clínica (Chen et al., 2012). Estudos mostram que o ML já permite avaliar riscos de doenças, otimizar planos de tratamento e reduzir readmissões hospitalares (Kruse et al., 2016). Intervenções personalizadas se tornaram viáveis graças à análise integrada de dados clínicos, genéticos e de estilo de vida (Nahar et al., 2024). Modelos treinados em grandes conjuntos de dados têm elevado a precisão diagnóstica em especialidades como radiologia e oncologia (Joy et al., 2024; Maraj et al., 2024; Rahman et al., 2024). Em casos específicos, algoritmos atingiram desempenho comparável ao de especialistas, como na classificação de lesões de pele (Youssef, 2014). Além disso, tais sistemas contribuem para a medicina personalizada ao prever a eficácia de tratamentos em nível individual (Lin et al., 2017).

Do ponto de vista administrativo, a integração de *Big Data* e ML possibilita maior eficiência operacional, melhor alocação de recursos e redução de custos (Istepanian & Alanzi, 2018). Também favorece a detecção precoce de riscos de saúde, como doenças crônicas, permitindo intervenções preventivas (Wan et al., 2017; Jensen et al., 2012).

Apesar dos benefícios, a adoção bem-sucedida dessas tecnologias enfrenta desafios relacionados à privacidade, interoperabilidade e qualidade dos dados. A fragmentação das informações em saúde exige padronização para garantir integração fluida entre sistemas (Wang et al., 2016), além de soluções seguras e interoperáveis que conciliem proteção de dados com uso clínico eficiente (Chen et al., 2016).

O *Big Data* e *Machine Learning* (ML) representa ferramentas transformadoras para a saúde contemporânea, capazes de impulsionar avanços em diagnóstico, personalização de tratamentos, gestão hospitalar e prevenção de doenças, embora ainda demandem superação de barreiras regulatórias e técnicas para sua plena consolidação.

2.1 Análise Preditiva na Saúde

A análise preditiva surgiu como uma ferramenta poderosa na área da saúde, oferecendo a capacidade de antecipar os resultados dos pacientes, otimizar tratamentos e reduzir custos por meio da tomada de decisões baseada em dados. A análise preditiva envolve o uso de Algoritmos Estatísticos, Modelos de *Machine Learning* (ML) e grandes Conjuntos de Dados para identificar padrões e prever eventos futuros (Qiu et al., 2016). Na área da saúde, os modelos preditivos são valiosos para prever a progressão da doença, readmissões hospitalares e fatores de risco dos pacientes, auxiliando, em última análise, os médicos a tomar decisões informadas. Archana e Anita (2015) argumentam que a análise preditiva permite que os profissionais de saúde antecipem as condições médicas antes que elas piorem, proporcionando oportunidades para intervenções mais precoces e melhores resultados para os pacientes. Essa crescente relevância da análise preditiva é impulsionada pela crescente disponibilidade de dados de saúde provenientes de prontuários eletrônicos de saúde (PEs), dispositivos vestíveis e dados genômicos, que podem ser aproveitados para refinar as previsões e melhorar a prestação geral de serviços de saúde.

A análise preditiva tem se destacado como ferramenta essencial para a prevenção e o gerenciamento de doenças, apoiada pelo uso de algoritmos de *Machine Learning* (ML) e técnicas de Deep Learning. Chu et al. (2015) demonstraram a eficácia de modelos preditivos baseados em prontuários eletrônicos (EHR) para prever mortalidade, tempo de internação e necessidade de readmissão, enquanto Chen et al. (2016) evidenciaram o potencial desses algoritmos para detectar precocemente casos de sepse em pacientes hospitalizados. Esses avanços ressaltam a capacidade da análise preditiva em antecipar diagnósticos, prevenir complicações e melhorar os resultados clínicos.

Casos de sucesso em instituições renomadas reforçam a aplicabilidade da tecnologia. A Cleveland Clinic, por exemplo, implementou modelos para avaliar riscos cardiovasculares, permitindo tratamentos personalizados (Oliver et al., 2004). A Kaiser Permanente utilizou a análise preditiva para identificar pacientes de alto risco e reduzir readmissões hospitalares por

meio de intervenções específicas (Herland et al., 2014). Já o Hospital Mount Sinai, em Nova York, aplicou modelos para prever surtos e monitorar a propagação de doenças infecciosas, contribuindo para respostas mais rápidas da saúde pública (White, 2014).

Apesar dos benefícios, desafios persistem. A qualidade e disponibilidade dos dados representam obstáculos, já que os registros de saúde são frequentemente fragmentados e inconsistentes, afetando a precisão das previsões (Adhikari et al., 2024; Qian et al., 2014; Tamal et al., 2024; Thapa et al., 2024). Questões de privacidade e segurança também são críticas, exigindo conformidade com regulamentações como a HIPAA (Kumar et al., 2015). Além disso, a integração da análise preditiva nos fluxos clínicos enfrenta barreiras técnicas e culturais, sendo necessária infraestrutura robusta, adesão organizacional e avaliações contínuas (Arora & Sharma, 2012).

Outro aspecto relevante são os desafios éticos, sobretudo o risco de viés algorítmico. Wang et al. (2016) alertam que modelos podem reproduzir desigualdades existentes nos dados de treinamento, comprometendo a equidade no atendimento. Para mitigar tais riscos, é imprescindível utilizar conjuntos de dados diversos e monitorar continuamente a imparcialidade dos modelos.

Entretanto mesmo que enfrente barreiras técnicas, regulatórias e éticas, a análise preditiva possui um papel transformador no cuidado em saúde. A adoção responsável dessa tecnologia pode ampliar sua contribuição para diagnósticos precoces, personalização de tratamentos, eficiência operacional e justiça social no acesso aos cuidados (Jim et al., 2024).

2.2 Plataformas de gerenciamento de dados na área da saúde

As plataformas de dados em saúde, como os registros eletrônicos (EHRs) e soluções em nuvem, tornaram-se essenciais para o gerenciamento e a disseminação de informações entre instituições médicas. Os EHRs possibilitam a coleta, armazenamento e compartilhamento sistemático de dados, configurando-se como pilares da gestão de saúde (Yin & Schütze, 2015). As soluções em nuvem ampliaram a escalabilidade e acessibilidade dos dados, permitindo o processamento remoto de grandes volumes e oferecendo opções econômicas para lidar com a crescente demanda informacional (Marcoon et al., 2013; Zhang et al., 2017).

A incorporação de *Big Data* a essas plataformas tem impulsionado avanços clínicos e administrativos. Sua aplicação em EHRs, por exemplo, viabiliza análises preditivas para antecipar resultados de pacientes e identificar riscos à saúde, favorecendo intervenções precoces

e atendimento personalizado (Youssef, 2014; Zhang et al., 2016). Contudo, a fragmentação dos dados, oriundos de múltiplas fontes — como registros clínicos, exames de imagem e dispositivos vestíveis —, representa desafio crítico para integração e compatibilidade (Zhang et al., 2016).

O debate entre plataformas centralizadas e descentralizadas expõe benefícios e limitações de cada abordagem. Sistemas centralizados, como os EHRs em nuvem, favorecem acessibilidade, continuidade do cuidado e recursos robustos de segurança (Picciano, 2012; Mounia & Habiba, 2015), mas enfrentam riscos de falhas únicas e violações de dados (Lin et al., 2016). Em contrapartida, sistemas descentralizados, como os baseados em blockchain, oferecem maior segurança, transparência e interoperabilidade entre diferentes entidades de saúde (Lusher et al., 2013). Ainda assim, sua implementação esbarra em complexidades técnicas e dificuldades de integração com infraestruturas já existentes.

A evolução das plataformas de dados de saúde depende do equilíbrio entre acessibilidade, segurança e interoperabilidade. A integração efetiva de *Big Data*, aliada a modelos inovadores de arquitetura centralizada ou descentralizada, será decisiva para transformar a gestão de dados e otimizar os resultados clínicos e administrativos nos sistemas de saúde contemporâneos.

2.3 *Machine Learning* (ML) na previsão de doenças e avaliação de risco

O *Machine Learning* (ML) consolidou-se como ferramenta essencial na saúde para modelagem preditiva de surtos e avaliação de riscos de pacientes, apoiando decisões clínicas baseadas em dados. Algoritmos são capazes de processar grandes volumes de informações, revelando padrões invisíveis à análise humana. Por exemplo, modelos têm sido aplicados para prever a disseminação de doenças infecciosas, como gripe e COVID-19, a partir de dados ambientais, de viagens e mobilidade populacional (Chen et al., 2017). Em paralelo, Chen e Zhang (2014) demonstraram a capacidade do ML de prever riscos de mortalidade por meio da análise de prontuários eletrônicos (PEs), permitindo melhor alocação de recursos e intervenções oportunas. No contexto hospitalar, estudos também evidenciaram seu uso na detecção precoce da deterioração clínica em pacientes de terapia intensiva (Lusher et al., 2013).

Casos específicos reforçam a eficácia da tecnologia em diferentes condições médicas. Modelos já foram empregados na previsão de fibrilação atrial a partir de eletrocardiogramas (K, 2015), na estimativa de readmissões e tempo de internação por meio de aprendizado

profundo (Rajkumar et al., 2018) e na oncologia, para identificar riscos de recidiva do câncer com base em dados genéticos e terapêuticos (Lusher et al., 2013). Esses exemplos evidenciam a versatilidade do ML, capaz de melhorar diagnósticos, prever eventos adversos e apoiar decisões clínicas.

Outro campo de destaque é a medicina personalizada. O ML pode integrar informações genéticas, histórico clínico e fatores de estilo de vida para recomendar intervenções específicas. K (2015) apontou que esses modelos já são aplicados no manejo de doenças crônicas, como diabetes e cardiovasculares. Na oncologia, Sagioglu e Sinanc (2013) destacaram sua eficácia em prever respostas individuais à quimioterapia e radioterapia.

O *Machine Learning* (ML) oferece contribuições significativas para diagnóstico, previsão de riscos e personalização de tratamentos. No entanto, sua plena integração clínica demanda maior transparência, diversidade de dados e aprimoramento contínuo dos modelos, a fim de garantir efetividade e justiça no cuidado em saúde.

3 PRINCÍPIOS DA REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL EM SAÚDE

A transparência é o princípio ético mais citado nas diretrizes de IA e significa fornecer informações suficientes sobre como as tecnologias funcionam antes de sua implementação. O mecanismo principal para concretizá-la é o direito à explicação de decisões automatizadas, inspirado no Regulamento Europeu de Proteção de Dados (GDPR). Esse direito garante ao indivíduo não apenas uma explicação inteligível sobre como a decisão foi tomada, mas também a possibilidade de a questionar e solicitar revisão.

No setor de saúde, esse direito é fundamental, pois decisões automatizadas já são capazes, por exemplo, de definir critérios para transplantes de órgãos, o que pode impactar diretamente listas de espera. Portanto, decisões críticas não devem se basear unicamente em sistemas de “caixa-preta”.

A LGPD garante aos titulares de dados o direito de solicitar a revisão de decisões automatizadas, mas diferentemente do GDPR, não prevê a obrigatoriedade de intervenção humana. O direito à explicação não está explicitamente na lei, mas decorre de sua interpretação sistemática junto à Constituição e ao Código de Defesa do Consumidor.

A aplicação do direito à explicação na saúde envolve desafios específicos. Cabe às autoridades regulatórias, como a ANPD (Autoridade Nacional de Proteção de Dados), a ANVISA e conselhos profissionais, definir os limites e mecanismos para que esse direito seja viabilizado.

4 DISCUSSÃO

O *Machine Learning* (ML) tornou-se uma ferramenta essencial na área da saúde para modelagem preditiva de surtos de doenças e avaliação de riscos para pacientes, permitindo que os profissionais de saúde tomem decisões baseadas em dados. Algoritmos de ML podem processar grandes quantidades de dados de saúde, identificando padrões que podem não ser imediatamente aparentes para médicos humanos. Por exemplo, modelos têm sido usados para prever a disseminação de doenças infecciosas, como gripe e COVID-19, por meio da análise de dados ambientais, padrões de viagem e movimentos populacionais (Chen et al., 2017). Da mesma forma, Chen e Zhang (2014) demonstraram como Modelos de *Machine Learning* (ML) podem prever o risco de mortalidade de pacientes por meio da análise de prontuários eletrônicos de saúde (PEs). Esses modelos preditivos permitem que os sistemas de saúde aloquem recursos de forma mais eficaz, garantam intervenções oportunas e potencialmente previnam surtos de doenças. Além disso, estudos como o de Lusher et al. (2013) mostraram como o *Machine Learning* (ML) pode prever a deterioração de pacientes em ambientes de terapia intensiva, melhorando os resultados ao identificar pacientes de alto risco que podem exigir monitoramento mais intensivo ou tratamento precoce. Vários estudos de caso destacaram a eficácia do ML na identificação de pacientes de alto risco em diversas condições médicas.

Da mesma forma, Rajkumar et al. (2018) usaram modelos de aprendizado profundo para prever com precisão os resultados dos pacientes, como readmissões e tempo de internação hospitalar, com base em dados de prontuários eletrônicos. Outro exemplo notável é a aplicação de ML em oncologia, onde modelos são usados para identificar pacientes com câncer com maior risco de recidiva com base em dados genéticos e de tratamento (Lusher et al., 2013). Esses estudos de caso ilustram a versatilidade do *Machine Learning* (ML) em vários domínios da área da saúde, destacando sua capacidade de melhorar a precisão do diagnóstico, prever eventos adversos e aprimorar a tomada de decisões clínicas. Uma de suas vantagens mais significativas é a capacidade do *Machine Learning* (ML) de oferecer suporte a planos de tratamento personalizados. Ao analisar dados específicos do paciente, incluindo informações genéticas, fatores de estilo de vida e registros históricos de saúde, os modelos de *Machine Learning* (ML) podem prever o tratamento mais eficaz para cada paciente. K (2015) relatou que algoritmos do *Machine Learning* (ML) usados em medicina de precisão poderiam adaptar tratamentos para condições crônicas como diabetes e doenças cardiovasculares. Além disso, Sagioglu e Sinanc (2013) descobriram que os modelos do *Machine Learning* (ML) eram

particularmente eficazes em oncologia, onde podiam prever as respostas dos pacientes à quimioterapia e à radioterapia com base em perfis genéticos individuais. No entanto, apesar desses avanços, a precisão preditiva dos modelos de *Machine Learning* (ML) não é limitada. Qiu e Sha (2009) destacaram potenciais preocupações éticas, incluindo viés algorítmico, que pode levar a resultados de tratamento desiguais para diferentes grupos demográficos. Além disso, os modelos de *Machine Learning* (ML) são frequentemente percebidos como "caixas-pretas", dificultando a compreensão dos médicos sobre como as previsões são feitas, o que pode reduzir a confiança na tecnologia (Bandyopadhyay et al., 2014). Essas limitações enfatizam a necessidade de refinamento contínuo dos modelos de *Machine Learning* (ML) e a inclusão de conjuntos de dados mais diversos para garantir resultados equitativos na área da saúde.

5 CONCLUSÃO

Os resultados deste estudo destacam o potencial transformador do *Big Data* e do *Machine Learning* (ML) na reformulação da saúde, particularmente na melhoria da análise preditiva, da medicina personalizada e da eficiência operacional. Nota-se que a adaptabilidade em tempo real dos modelos de *Machine Learning* (ML) é fundamental para gerenciar os riscos dos pacientes de forma mais eficiente do que os sistemas anteriores, que frequentemente demoravam a se adaptar às mudanças nas condições de saúde.

Em termos de medicina personalizada, o *Machine Learning* (ML) provou ser um divisor de águas na adaptação de tratamentos a pacientes individuais, oferecendo uma vantagem distinta em relação às abordagens tradicionais e universais. A capacidade dos modelos de *Machine Learning* (ML) de analisar grandes conjuntos de dados, incluindo fatores genéticos e de estilo de vida, permite intervenções mais direcionadas, particularmente no manejo de doenças crônicas.

Por fim, considerações éticas permanecem em primeiro plano na discussão sobre o uso de *Big Data* e *Machine Learning* (ML) na área da saúde. A falta de transparência contrasta com os métodos diagnósticos anteriores, em que os processos de tomada de decisão eram mais transparentes e acessíveis para os médicos explicarem aos pacientes. Além disso, ainda existem as preocupações sobre o viés algorítmico, observando que os sistemas de ML, se treinados em conjuntos de dados tendenciosos, podem perpetuar disparidades no acesso e tratamento da saúde.

Ao reunir avaliações críticas e práticas de diferentes métodos de *Machine Learning* (ML), este trabalho oferece subsídios para pesquisadores e profissionais compreenderem as

possibilidades e os desafios de sua implementação. Conclui-se, portanto, que a integração responsável de *Big Data* e *Machine Learning* (ML) representa um caminho promissor para apoiar processos de tomada de decisão mais eficazes, contribuindo tanto para a melhoria do cuidado em saúde quanto para a evolução das instituições hospitalares.

Desta forma, este artigo cumpriu o objetivo de examinar a aplicação do *Machine Learning* (ML) no contexto do *Big Data* na área da saúde, evidenciando como essas tecnologias têm potencial para transformar a prática clínica e a gestão hospitalar. A análise apresentada destacou não apenas os avanços obtidos em diagnósticos, medicina personalizada, prevenção e eficiência operacional, mas também as limitações relacionadas à privacidade, interoperabilidade, qualidade dos dados e questões éticas.

REFERÊNCIAS

- BANDYOPADHYAY, S. et al. Data mining for censored time-to-event data: a Bayesian network model for predicting cardiovascular risk from electronic health record data. *Data Mining and Knowledge Discovery*, v. 29, n. 4, p. 1033-1069, 2014. Disponível em: <https://doi.org/10.1007/s10618-014-0386-6>. Acesso em: 24 set. 2025.
- CHEN, H.; YU, Z.; KANG, W.; ZHANG, B.; LI, H.; YANG, B.; LIU, D. Evolving support vector machines using fruit fly optimization for medical data classification. *Knowledge-Based Systems*, v. 96, p. 61–75, 2016.
- HUANG, Z.; DESTECH, P. Research on the innovation of e-business talents training mode under the background of big data. In: INTERNATIONAL CONFERENCE ON E-COMMERCE AND ECONOMIC DEVELOPMENT, 2018. *DEStech Transactions on Business, Economics and Management*. Anais [...]. 2018. p. 48–52.
- KUMAR, N. M. S. et al. Predictive Methodology for Diabetic Data Analysis in Big Data. *Procedia Computer Science*, v. 50, p. 203-208, 2015. Disponível em: <https://doi.org/10.1016/j.procs.2015.04.069>. Acesso em: 24 set. 2025.
- LIN, K. et al. Enhanced Fingerprinting and Trajectory Prediction for IoT Localization in Smart Buildings. *IEEE Transactions on Automation Science and Engineering*, v. 13, n. 3, p. 1294-1307, 2016. Disponível em: <https://doi.org/10.1109/tase.2016.2543242>. Acesso em: 24 set. 2025.
- LIN, K. et al. Localization Based on Social Big Data Analysis in the Vehicular Networks. *IEEE Transactions on Industrial Informatics*, v. 13, n. 4, p. 1932-1940, 2017. Disponível em: <https://doi.org/10.1109/tii.2016.2641467>. Acesso em: 24 set. 2025.

- LUSHER, S. J. et al. Data-driven medicinal chemistry in the era of big data. *Drug Discovery Today*, v. 19, n. 7, p. 859-868, 2013. Disponível em: <https://doi.org/10.1016/j.drudis.2013.12.004>. Acesso em: 24 set. 2025.
- MAHADEVKAR, S. et al. A review on machine learning styles in computer vision—techniques and future directions. *IEEE Access*, v. 10, p. 107293–107329, 2022.
- MARCOON, S. et al. HEART score to further risk stratify patients with low TIMI scores. *Critical Pathways in Cardiology*, v. 12, n. 1, p. 1-5, 2013. Disponível em: <https://doi.org/10.1097/hpc.0b013e31827377e1>. Acesso em: 24 set. 2025.
- MARAJ, M. A. H. S. I.; MAHMUD, N. U. Information Systems in Health Management: Innovations And Challenges In The Digital Era. *International Journal of Health and Medical*, v. 1, n. 2, p. 14-25, 2024. Disponível em: <https://doi.org/10.62304/ijhm.v1i2.128>. Acesso em: 24 set. 2025.
- MOUNIA, B.; HABIBA, C. EUSPN/ICTH - Big Data Privacy in Healthcare Moroccan Context. *Procedia Computer Science*, v. 63, p. 575-580, 2015. Disponível em: <https://doi.org/10.1016/j.procs.2015.08.387>. Acesso em: 24 set. 2025.
- NAHAR, J. et al. Advanced Predictive Analytics For Comprehensive Risk Assessment In Financial Markets: Strategic Applications And Sector-Wide Implications. *Global Mainstream Journal of Business, Economics, Development & Project Management*, v. 3, n. 4, p. 39-53, 2024. Disponível em: <https://doi.org/10.62304/jbedpm.v3i4.148>. Acesso em: 24 set. 2025.
- TAHA, K.; SALAH, K.; YOO, P. Clustering the dominant defective patterns in wafer maps. *IEEE Transactions on Semiconductor Manufacturing*, v. 31, n. 1, p. 156–165, 2025.
- WANG, S. et al. Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression. *IEEE Access*, v. 4, p. 7567-7576, 2016. Disponível em: <https://doi.org/10.1109/access.2016.2620996>. Acesso em: 24 set. 2025.
- WHITE, S. E. A review of big data in health care: challenges and opportunities. *Open Access Bioinformatics*, v. 6, p. 13-18, 2014. Disponível em: <https://doi.org/10.2147/oab.s50519>. Acesso em: 24 set. 2025.
- YOUSSEF, A. E. A Framework for Secure Healthcare Systems Based on Big Data Analytics in Mobile Cloud Computing Environments. *The International Journal of Ambient Systems and Applications*, v. 2, n. 2, p. 1-11, 2014. Disponível em: <https://doi.org/10.5121/ijasa.2014.2201>. Acesso em: 24 set. 2025.
- ZHANG, Y. et al. Fractal Dimension Estimation for Developing Pathological Brain Detection System Based on Minkowski-Bouligand Method. *IEEE Access*, v. 4, p.