



## INTRODUÇÃO AO ESTUDO DE GERAÇÃO DE DADOS SINTÉTICOS COMO FERRAMENTA PARA O PROCESSO DE ENSINO- APRENDIZAGEM EM ENGENHARIA.

DOI: 10.37702/2175-957X.COBENGE.2024.5114

**Autores:** JOÃO PAULO DE OLIVEIRA MACEDO, GILMAR BARRETO, PAULO VICTOR DE OLIVEIRA MIGUEL

**Resumo:** A humanidade está vivendo um momento histórico único, pois estamos lidando com uma enorme quantidade de dados, conhecida como Big Data. Há uma tendência crescente de coletar informações, analisá-las e tirar conclusões, além de desenvolver algoritmos e sistemas baseados em Inteligência Artificial (IA) cada vez mais avançados, que podem ser aplicados em diversas áreas. No entanto, essa necessidade de sistemas especializados e inteligentes traz consigo preocupações relacionadas à privacidade, integridade e escassez de dados. Este trabalho destaca a importância dos dados como base para estudantes, desenvolvedores e pesquisadores na modelagem de sistemas e solução de problemas. No entanto, em muitas situações, a extração de dados reais do mundo se torna um processo lento e inviável devido ao tempo necessário para monitoramento e coleta, interferências externas ou viés dos dados coletados. Isso pode tornar algumas aplicações inviáveis com dados puramente reais. É nesse contexto que se faz necessária a adoção de novas técnicas de geração e modelagem de dados na Engenharia. A geração de dados sintéticos é uma maneira de garantir a privacidade, integridade e volume de informações, permitindo controlar o tempo, a quantidade, a integridade e a privacidade dos dados necessários para treinar sistemas inteligentes. Além disso, este trabalho aborda os fundamentos da relação entre professores e alunos, demonstrando como, ao longo do tempo, a Engenharia tem impactado significativamente o desenvolvimento da sociedade, estabelecendo um novo padrão técnico e científico. A união entre ciência e tecnologia na Engenharia é responsável por renovar as esperanças e alimentar as expectativas sociais para o futuro, com foco na superação de diversos problemas. A sociedade moderna depende da Engenharia, e os avanços tecnológicos são resultados desse trabalho. Portanto, é de fundamental importância incorporar formas eficazes de desenvolvimento de competências e habilidades nos alunos, de modo que eles possam atender às necessidades sociais com uma base científica sólida e uma compreensão global dos problemas mais variados. Portanto, neste trabalho, propomos a incorporação da geração de dados sintéticos por meio de uma Rede Gerativa Adversária (GAN) como ferramenta para o processo educacional. Essa abordagem está vinculada a uma proposta metodológica de ensino, na qual o professor desempenha o papel de orientador e o aluno, o papel de protagonista em seu próprio processo de aprendizagem.

**Palavras-chave:** dados sintéticos, processo de ensino-aprendizagem, inteligência artificial.

# INTRODUÇÃO AO ESTUDO DE GERAÇÃO DE DADOS SINTÉTICOS COMO FERRAMENTA PARA O PROCESSO DE ENSINO- APRENDIZAGEM EM ENGENHARIA.

## 1 INTRODUÇÃO

A humanidade vive um momento singular em sua história, experimentando uma gama gigantesca de dados (*Big Data*), como tendência de reunir informações, analisá-las e chegar a conclusões, bem como, em desenvolver algoritmos e sistemas baseados em Inteligência Artificial (IA) cada dia mais robustos para as mais variadas aplicações.

A crescente necessidade de Sistemas Especialistas e Inteligentes, demandam a extração, análise e aplicação de dados, onde por vezes preocupações podem surgir envolvendo a privacidade, integridade e falta de dados. Os dados têm por objetivo servirem de base para que alunos, desenvolvedores e pesquisadores possam modelar sistemas e resolver problemas. Em diversas situações a sua extração do mundo real torna-se lenta e impraticável, seja pelo tempo necessário no monitoramento e coleta, interferências externas ou mesmo pelo viés que cada dado traz. Desta forma, com dados puramente reais algumas aplicações podem se tornar inviáveis.

É nesse sentido que a incorporação de novas técnicas de geração e modelagem de dados se faz necessária. A geração de dados sintéticos é uma maneira de garantir a privacidade, integridade e volume de informações, viabilizando o tempo, quantidade, integridade e privacidade do volume de dados necessários para o treinamento de Sistemas Inteligentes.

Essa nova maneira de geração de dados abre caminho para uma mudança radical na maneira como o desenvolvimento da Inteligência Artificial e dos demais assuntos dentro da área de Engenharia se faz nos dias de hoje e se consolidará no futuro, viabilizando a convergência de temas transversais, como relacionamento interpessoal, saúde, política, educação e ciência.

Além da maneira de ampliar a produção de dados, tornando-os menos vulneráveis a riscos cibernéticos relacionados à segurança e privacidade, é possível democratizar a quantidade de informação, favorecendo alunos, pesquisadores, pequenas empresas de desenvolvimento ou mesmo desenvolvedores individuais, para que qualquer pessoa possa competir com grandes organizações (geralmente monopolizadoras de um vasto banco de dados, alimentado por anos, pelas informações dos seus usuários). Essa quebra de paradigmas pode resultar em uma disrupção da IA e uma melhora significativa no processo de ensino-aprendizagem dentro dos cursos de Engenharia, minimizando o monopólio de grandes organizações e estabelecendo um processo crescente de aumento de qualidade nos diversos estudos de sistemas baseados em dados.

A base para a Inteligência Artificial são os dados, estabelecer processos matemáticos e computacionais que criem esses parâmetros sem a necessidade da intervenção humana para coletá-los e tratá-los, reflete um ganho significativo para a construção de qualquer modelo de *Machine Learning*. Atualmente a IA está presente nos mais variados campos do conhecimento e, principalmente, como suporte aos conteúdos estudados nas áreas de Engenharia, auxiliando com cálculos matemáticos, geração de imagens para realidade aumentada, auxílio na produção de apresentações, geração de ideias na descoberta de novos materiais e produtos, entre diversos outros. Neste contexto, o mundo real gera *insights* que os dados sintéticos nunca fomentarão, por outro lado, eles

segmentam de maneira surpreendente os limites necessários para treinamento e implementação de modelos computacionais assertivos para qualquer aplicação. A orientação do universo amostral dos dados sintéticos reflete o conjunto de informações necessárias para o modelo, reduzindo distorções externas desnecessárias e valorizando os aspectos realmente relevantes. Em algumas ocasiões é interessante a mescla entre dados sintéticos e dados reais.

Existem alguns desafios a serem vencidos com a geração de dados sintéticos, como por exemplo, a chamada adaptação do domínio, onde o modelo de Inteligência Artificial é treinado com um conjunto de dados sintéticos, mas é aplicado a um domínio do mundo real, o que pode resultar em divergências das respostas esperadas.

Dentro desta abordagem, este trabalho busca mostrar que a geração de dados sintéticos pode ser um componente importante do processo educacional, propondo uma metodologia a ser empregada para a incorporação desta temática como ferramenta de ensino-aprendizagem nos cursos de Engenharia.

A Engenharia ao longo do tempo impactou significativamente o desenvolvimento humano, adotando métodos relacionais e comparativos (baseados em dados) que alavancaram saltos tecnológicos e inovadores, superando problemas, metodologias e falta de recursos para subsidiar os avanços necessários nas mais variadas áreas do conhecimento. Sendo assim, encontrar ferramentas educacionais que viabilizem a inserção do aluno como protagonista do seu processo de aprendizagem é fundamental para sua formação, bem como para a disseminação do conhecimento gerado ao longo do tempo dentro dos grandes centros de pesquisas e que necessitam de fluidez para o mercado. É necessário estabelecer um ponto ótimo entre métodos matemáticos, recursos computacionais, metodologias de ensino e os mais novos avanços tecnológicos na área de Engenharia, para formar alunos engajados, preparados e dedicados em seu segmento de estudo, que supram as demandas de pesquisa e do mercado de trabalho.

Por essa óptica, a maneira como é abordada a formação educacional em Engenharia deve refletir a demanda imposta pela sociedade e, ao mesmo tempo, suprir a carência de pessoas qualificadas para resolução de problemas complexos dentro de sua área de estudo.

## 2 DADOS SINTÉTICOS E O PROCESSO DE ENSINO-APRENDIZAGEM

A Engenharia está apoiada em dados, estudar maneiras de gerá-los sinteticamente abre caminho para ampliação de recursos, inferência sobre estatísticas válidas, além de salientar informações importantes e minimizar a preocupação de ferir limites de confidencialidade e privacidade (RAGHUNATHAN, 2021).

Atualmente a sociedade é dependente de informações, quanto mais adequadas, passando por diversas perspectivas e procedimentos analíticos, maior o grau de credibilidade e precisão para sua utilização na tomada de decisões. Um exemplo deste fato é a adoção de políticas públicas assertivas na pandemia de Covid-19 (RAGHUNATHAN, 2021).

Na luta contra a pandemia de Covid-19, que teve seu início na cidade de Wuhan, na China, no ano de 2019, se espalhando por todos os continentes do mundo, com exceção da Antártida, motivou a comunidade científica a procurar soluções baseadas em sistemas computacionais para prevenção, diagnóstico e estimativas de informações. Alguns dos principais aspectos de pesquisas neste sentido, teve por objetivo a utilização de dados disponíveis para análise estatística e adoção da Inteligência Artificial como ferramenta de interpolação de parâmetros (SHUJA *et al.*, 2020).

A falta de dados e o não compartilhamento de informações entre entidades públicas e privadas, poderiam ter acarretado em um atraso significativo na contenção e tratamento da Covid-19. Desta forma, fica evidente que a manipulação de dados é fundamental para a resolução de problemas do mundo real, gerando ganhos significativos para a humanidade em diversas frentes.

Ainda que exista uma crescente necessidade de dados para os mais variados problemas, não pode ser deixada de lado a preocupação com a privacidade e a confidencialidade das informações, os indivíduos devem ser protegidos para que detalhes específicos de suas vidas, que compõem o banco de dados, não sejam acessados inadvertidamente por intrusos. A incorporação de dados de fontes distintas, envolvendo por exemplo, sensores, redes sociais e outros tipos de instrumentos de medição, são frutos do crescente uso do ambiente virtual, resultando da coleta de dados do mundo real, do espaço físico e do meio social, e dados do ambiente cibernético, onde, como fruto deste sistema, o próprio ser humano pode ser enquadrado como gerador ou consumidor de dados. Como forma de proteger a privacidade dos dados, existem pesquisas que buscam na criptografia a solução para esse problema, mesmo com indicadores que apontam para o alto custo computacional nesse processo (GATI *et al.*, 2021).

Há décadas existe a crescente necessidade de acompanhamento da geração de dados estatísticos e com os bancos de dados, como fruto desta lacuna brotou o ramo de engenharia de segurança de dados (ou engenharia e segurança da informação), visando incorporar sistemas da informação e métodos matemáticos para contingenciar a propagação indevida de dados, estabelecendo uma “transformação de privacidade” através de técnicas de codificação. São inúmeras as áreas contempladas, envolvendo: sistemas de informação para justiça criminal, sistemas de créditos bancários, sistemas de informação em saúde, sistemas de informação na área educacional etc. (DALENIUS, 1977).

Na área educacional, mais especificamente nos cursos de Engenharia, os dados permeiam todos os desafios a serem enfrentados. Na Engenharia Elétrica, por exemplo, as fontes de energias renováveis estão cada vez mais presentes para atender a demanda energética e reduzir os gases de efeito estufa, mas algumas dessas fontes são instáveis quanto ao ciclo de geração, o que necessita de estudos para integrar as formas de geração de energia, melhorar a confiabilidade da matriz energética e aumentar a eficiência de todo o sistema. Outros aspectos tangenciais podem ser levantados para correlacionar e validar modelos que justifiquem a aplicação ou não de fontes renováveis de energia, como: aportes financeiros, qualidade ambiental (impactos, degradação e preservação), desenvolvimento industrial, relação entre crescimento econômico e a matriz energética empregada, entre diversos outros fatores que poderiam ser citados como norteadores para pesquisas neste sentido (NUÑA *et al.*, 2024).

A Engenharia, de modo geral, tem a missão de encontrar soluções para problemas técnicos e aplicá-los em benefícios do ser humano. Para atingir esse objetivo, são utilizados conhecimentos das ciências humanas, exatas e naturais, levando em consideração recursos materiais, tecnológicos e econômicos, além de se amparar nas restrições legais, ambientais e outras impostas pelo próprio ser humano.

## 2.1 Caracterização dos dados

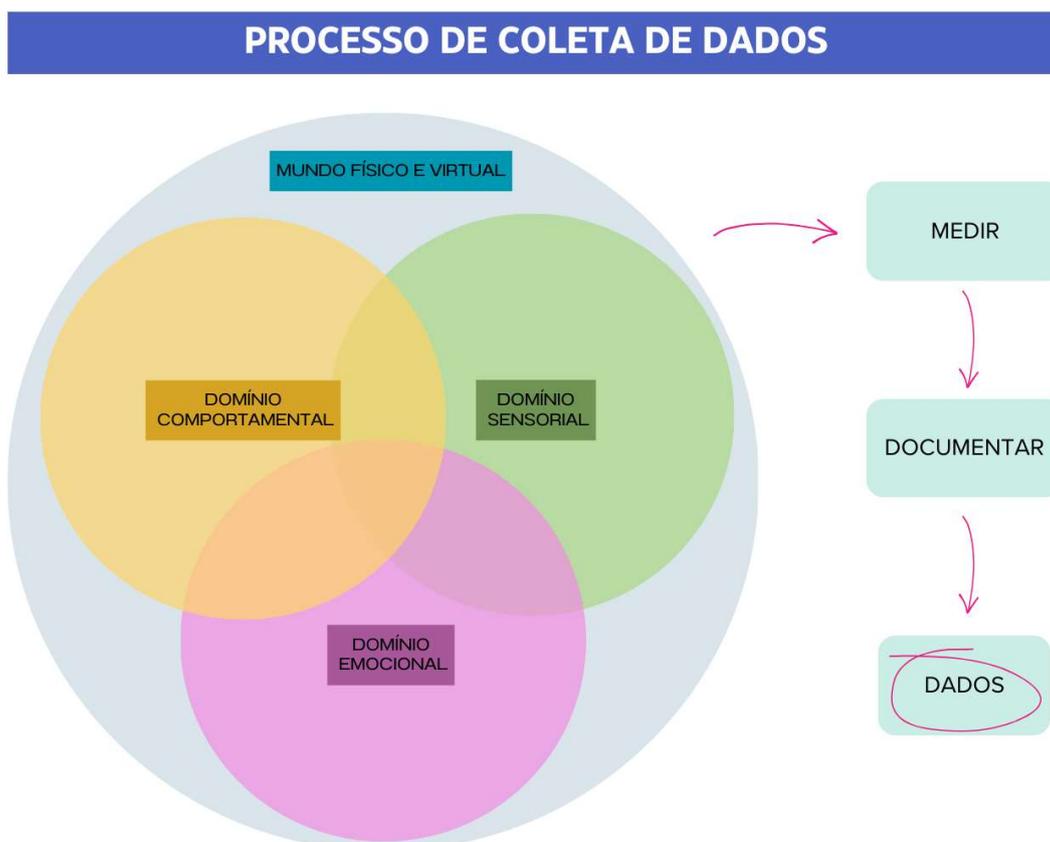
Os dados são frutos da documentação de medições realizadas no universo de estudo. Este universo pode ser caracterizado pelo conjunto de interações entre o agente e a ação realizada (PINHEIRO, 2024).

Dentro desta ideia, pode ser caracterizado um único universo, composto homogeneamente por duas partes onde a medição dos dados ocorrem, o Mundo Físico, que conhecemos e interagimos todos os dias, e o Mundo Virtual, o qual a cada dia estamos

mais imersos. Neste universo de discurso podem ser coletadas medições que refletem aspectos de três subconjuntos: Domínio comportamental, Domínio sensorial e Domínio emocional (figura 1).

No Domínio comportamental são observados aspectos relativos à maneira como o agente se comporta em determinadas situações. Em uma sala de aula, por exemplo, pode ser observado como os alunos reagem a um exercício: de maneira positiva, neutra ou negativa, sendo que: o fato decorre no Mundo Físico, o que se pretende medir é a atitude dos alunos, o instrumento de medição é a observação e a medida a ser documentada são as atitudes definidas em positivas, neutras ou negativas. Para o Domínio sensorial, até o momento, são coletadas informações do Mundo Físico, mas em um futuro próximo podem ser resultados de observações do Mundo Virtual, através do agente do processo. Para definir este domínio, tomemos como exemplo o ato de provar um alimento, o que se busca medir: aroma, o instrumento de medição: são os provadores e a medida: ruim, regular, bom ou excelente. Dentro do Domínio emocional as relações tornam-se nebulosas, dadas as dificuldades em estabelecer correlações entre as partes avaliadas. Mas ainda assim, é possível definir se o sentimento estudado pertence ao Mundo Físico e Virtual, o que se deseja medir, o instrumento a ser empregado e qual a medida (PINHEIRO, 2024).

Figura 1 - Fluxo para determinação de dados.



Fonte: Desenvolvido pelos autores.

Dentro de uma base de dados é definida a unidade observacional e vinculada a ela temos as variáveis do processo (ANDERSON; FINN, 2012).

## 2.2 Geração de dados sintéticos e exemplo de aplicação.

A utilização de máquinas de aprendizado, chamadas de Redes Adversárias Generativas (GAN's) do inglês *Generative Adversarial Network*, é uma ótima técnica para geração de dados sintéticos. Elas são capazes de gerar padrões sintéticos (imagens, sons, vetores entre outros) dentro de um domínio de interesse praticamente indistinguíveis dos padrões reais. Isso é possível pois existe uma associação entre uma rede generativa, que é treinada em conjunto com uma rede neural adversária. A rede neural adversária tem por objetivo servir como um classificador binário, tendo que se especializar para diferenciar o que é real do que é sintético. As duas redes neurais, a generativa e a discriminativa, são adversárias, disputam entre si: a discriminatória busca cada vez mais melhorar sua distinção entre real e sintético, por outro lado, a rede generativa busca tornar a sua saída sintética menos discriminável. Ao final do processo de treinamento, atinge-se um equilíbrio, onde a rede classificadora não é capaz de distinguir entre a fonte real e a fonte sintética, desta forma, pode-se afirmar que a rede generativa aprendeu os padrões que se pretende gerar (GUI *et al.*, 2020).

Figura 2 - Rede GAN.



Fonte: Adaptada pelos autores de: <https://www.dca.fee.unicamp.br/~vonzuben/courses/>. Acesso: 10/04/2024

Essa técnica pode ser utilizada na geração de vetores em um universo n-dimensional para Interfaces Neurais Interativas (INI), simulando ondas cerebrais de percepções do indivíduo de qualquer um dos seus Domínios (MIGUEL, 2010).

Os dados reais são provenientes de uma base de dados reais. Já a rede generativa aprende a gerar dados sintéticos similares aos reais, estabelecendo o Equilíbrio de Nash entre a rede generativa e a discriminativa. Os dados provenientes da rede generativa, quando atingido o equilíbrio, são indistinguíveis dos reais, favorecendo a quantidade, qualidade e privacidade do volume de informações, criando cenários distintos para o modelo, tornando-os robustos e adaptáveis, além de possibilitar o teste e avaliação dentro de um ambiente seguro e hipotético, sem a necessidade de intervenção humana (VON ZUBEN; VIANNA, 2023). Com as Interfaces Neurais Interativas (INI) é possível estabelecer uma comunicação homem-máquina não invasiva, vencendo a uma antiga fronteira entre o ser humano e o computador, incorporando o Mundo Físico ao Mundo Virtual (MIGUEL, 2010).

### 2.3 Processo de Ensino-Aprendizagem

Ao longo do tempo, a Engenharia impactou consideravelmente a maneira como a sociedade se relaciona, estabelecendo um novo padrão técnico e científico. A união entre ciência e tecnologia na Engenharia, é a responsável por renovar esperanças e alimentar expectativas sociais para o futuro, focando em superar os mais variados problemas. Desta forma, fica evidenciado que a vida humana, tal como a conhecemos, está submetida aos mais variados aspectos originados da Engenharia. Tomando por base que a sociedade moderna é dependente da Engenharia e que os avanços tecnológicos são frutos deste trabalho, é de fundamental importância incorporar maneiras efetivas de desenvolver competências e habilidades nos estudantes, para que eles suplementem as necessidades sociais com uma base científica sólida, além de uma interpretação global dos mais variados problemas aos quais forem submetidos (VALÉRIO; BAZZO, 2006).

Neste sentido, a utilização de metodologias ativas proporciona uma interligação simbiótica entre dois mundos: o físico e o digital, quando amparada por recursos computacionais, introduzindo uma visão de sala de aula ampliada e induzindo o aluno a múltiplos espaços do cotidiano em que, professor e aluno, compartilham ideias de maneira clara e assertiva (MORÁN, 2015).

A escola atual é caracterizada como um tempo-espço que resiste às mudanças das gerações, onde os sujeitos-clientes são avaliados constantemente, mas os sujeitos-agentes não, ancorando-se em modelos metodológicos ultrapassados, existe algo de bom a ser preservado dentro de alguns aspectos do modelo de ensino atual, mas a velocidade de comunicação, de mudanças de cenários, entre outros aspectos, origina um descompasso entre o formato atual e a necessidade futura do ensino (VIEIRA, 2014).

É crucial que sejam estabelecidos modelos educacionais inovadores e abrangentes, aguçando nos alunos a vontade de buscarem novos conceitos, criarem habilidades e competências, refletindo as carências impostas pela sociedade. A inovação no ensino superior é composta por um conjunto de ações que afetam pontos-chaves, bem como eixos constitutivos da organização do ensino universitário, alavancadas por mudanças sociais ou por reflexões intrínsecas à missão e organização da estrutura educacional. Um dos grandes saltos, com a inserção de metodologias ativas nesse modelo, é o maior envolvimento do aluno com o curso e, conseqüentemente, um melhor aprendizado. A aprendizagem, quando embasada por auto iniciativa, alcança padrões afetivos e intelectuais mais profundos, tornando-se mais duradoura e sólida. Para atingir esse padrão é necessário que o professor esteja próximo ao aluno, como um tutor/orientador e não apenas como o ator principal, ministrando aulas expositivas descontextualizadas do cotidiano do aluno (MASETTO, 2004).

Estruturar um processo de ensino mais efetivo, que transcenda as tradicionais aulas mecanizadas, requer uma nova abordagem pelo docente, é necessária a quebra de diversas barreiras, forçando-o a uma reestruturação constante de postura teórica, prática e trato social (SKINNER, 2016)

Dentre as mais variadas ferramentas a serem utilizadas no processo de ensino-aprendizagem, que solucionam as dificuldades impostas pela evolução da educação em Engenharia, a incorporação de sistemas computacionais têm mostrado uma grande aceitação e resultados muito promissores entre os estudantes universitários. As novas gerações, principalmente os jovens nascidos na década de 1990 em diante, apresentam um desenvolvimento educacional muito melhor quando são auxiliados por recursos tecnológicos computacionais e ferramentas virtuais de aprendizagem (QUINTANILHA, 2017).

O processo de ensino-aprendizagem é cercado por um complexo conjunto de ações que vão além do conteúdo ministrado pelo professor. O processo de ensino é resultado de

uma relação entre o professor e o aluno. O sucesso do ensino provém do entusiasmo do professor à ciência e aos seus alunos, algo que pode ser canalizado mediante planejamento e metodologias, sobretudo para motivar os alunos pela realização, por vontade própria, movidos por seus esforços intelectuais e morais que a aprendizagem necessita. Neste sentido, as principais relações entre aluno, professor e o embasamento teórico da disciplina para obtenção de sucesso no processo de ensino podem ser elencados como: a) Aluno: motivação, conhecimentos prévios, relação com o professor e atitude com a disciplina; b) Assunto: estrutura: componentes e relações, tipos de aprendizagem, ordem de apresentação; c) Professor: situação estimuladora ambiental, comunicação verbal de instrução, informação ao aluno sobre seus progressos, relacionamento com o aluno e atitude com a matéria ensinada. Já a aprendizagem, é a manifestação da inteligência, algo que está intrinsecamente vinculado a processos neurológicos do corpo. A inteligência modifica a estrutura de aprendizagem e o próprio processo. A construção é baseada em interações neurológicas entre o indivíduo e o meio ambiente. Ou ainda, a aprendizagem pode ser encarada como consequência dos estímulos ambientais, focando que atividades que sejam recompensadas ou reforçadas e que produzam satisfação tenderá a ser repetida e aprendida pelo aluno (SANTOS, 2001).

O processo de ensino-aprendizagem, da forma como abordado até aqui, é definido por quatro pilares básicos: o professor, o aluno, o conteúdo e as variáveis ambientais (características da instituição de ensino), sendo que cada componente pode ter maior ou menor influência sobre o processo. Ao aluno cabe a capacidade, a velocidade de aprendizagem, conhecimentos prévios, disposição e boa vontade, interesse e aspectos relacionados. O conteúdo, por sua vez, vincula-se às dimensões do aluno, tendo que gerar significado e valor, além de aplicabilidade prática a ele. A instituição de ensino, estabelece o processo educacional, baseado nas crenças dos dirigentes e na liderança. O professor é quem estabelece as relações entre professor-aluno, aspectos cognitivos e técnicos-didáticos, atitudes do educador, capacidade inovadora e comprometimento com o processo de ensino-aprendizagem (MOREIRA, 1986).

## 2.4 Proposta metodológica

A inserção de novas metodologias de ensino, dentro da área de Engenharia, tem motivado a observação de ganhos significativos no processo de ensino-aprendizagem, onde o conteúdo não é o único fator decisivo para que o aluno atinja um alto nível de aprendizagem. Nesse sentido, propor novos métodos e entender como a articulação entre o professor, o conteúdo, o aluno e os recursos tecnológicos são aplicados é crucial para que o aluno atinja um bom desenvolvimento intelectual e pessoal.

O ensino está em constante transformação, bem como, o ato de aprender necessita de estímulos novos a cada geração. Distante de ser uma solução para todos os problemas enfrentados com falta de recursos financeiros para pesquisa, diferenças educacionais entre os alunos, falta de recursos físicos nas instituições de ensino, entre outros aspectos que prejudicam o trabalho do professor e, conseqüentemente, o desenvolvimento de habilidades e competências dos alunos, este trabalho apresenta uma proposta metodológica para mitigar esse problema, onde a interação entre o professor, o aluno, os conteúdos e os recursos tecnológicos são aproveitados com a implementação de dados sintéticos como uma ferramenta para o processo de ensino-aprendizagem (figura 3).

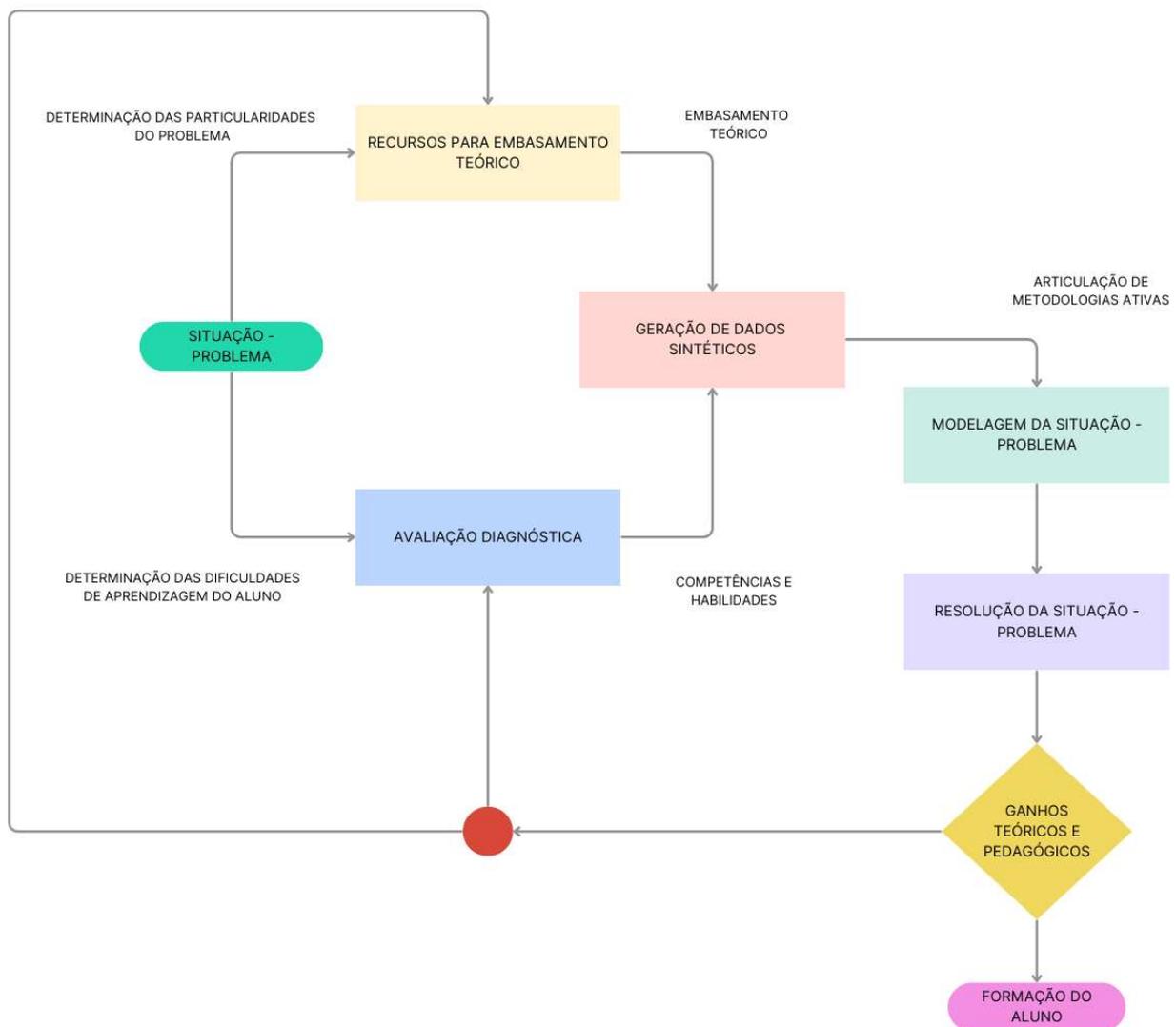
A proposta está baseada em etapas que são interligadas entre si, através de um processo realimentado, onde o professor é responsável pelo acompanhamento da evolução educacional do aluno e o aluno, por sua vez, é o agente do seu processo de aprendizagem.

O modelo inicia-se com uma situação-problema. O problema a ser solucionado dentro da Engenharia pode ser proposto diretamente pelo docente ou definido em conjunto

com a turma, é importante que o tema a ser trabalhado estimule os alunos, que seja algo relevante às suas expectativas e que gere ganhos concretos para a formação técnica e pessoal do estudante. O tema deve estar entre dois limiares para o aluno: não muito fácil, para que ele não perca logo o interesse no trabalho, nem muito difícil, para que não gere a sensação de incapacidade e frustração.

Com a situação-problema definida, o docente é responsável por um conjunto de ações dentro de uma subcategoria de avaliação técnica e pedagógica. Aqui são determinadas as particularidades do problema e, com isso, quais serão os recursos necessários para fornecer aos alunos o embasamento teórico na resolução do problema. Podem ser definidos: trabalhos em grupo, apresentação de vídeos, estudos dirigidos, caberá ao professor elencar quais os melhores recursos para essa finalidade.

Figura 3 - Proposta Metodológica.



Fonte: Desenvolvido pelos autores.

Paralelamente, o professor fará uma avaliação das dificuldades de aprendizagem do aluno, culminando com a formulação de uma avaliação diagnóstica. Com a avaliação diagnóstica e com a definição dos recursos didáticos necessários para o embasamento teórico, o docente tem capacidade de propor a incorporação da geração de dados sintéticos

para resolução do problema, observando quais serão as competências e habilidades alavancadas no processo de ensino.

Com os dados gerados, cabe a modelagem e resolução da situação-problema. Nessa etapa a articulação de Metodologias Ativas pode resultar em ganhos significativos para toda a turma, o professor, como tutor/orientador, pode estimular a troca de conhecimento entre os alunos, buscando uma homogeneização de informação e ao mesmo tempo realizando um processo de recuperação contínua.

Ao final da resolução da situação-problema o docente deve avaliar quais os ganhos para o estudante, quando os ganhos são positivos, apontando para uma boa formação de competências e habilidades, o docente entende o processo de aprendizagem como atingido. Caso não sejam observados ganhos significativos para o aluno, o docente pode revisitar o processo, observando a Avaliação diagnóstica e os Recursos para o embasamento teórico.

### 3 Considerações FINAIS

Com base neste trabalho, conclui-se que a adoção de novas técnicas educacionais, como a incorporação de dados sintéticos nos cursos de Engenharia, pode contribuir para a formação dos estudantes, promovendo uma maior interação entre professores e alunos. Essa abordagem proporciona uma formação teórica e comportamental sólida, que engloba tanto o domínio técnico (matemática, física e ciências exatas) quanto conhecimentos específicos da área de atuação. Além disso, desenvolve habilidades analíticas para solução de problemas complexos, pensamento crítico, raciocínio lógico e competências comportamentais fundamentais, como comunicação eficaz, trabalho em equipe, colaboração, criatividade, inovação, aprendizagem contínua e adaptabilidade.

Nesse sentido, a geração de dados sintéticos oferece soluções inovadoras para os desafios da educação em Engenharia, apresentando os seguintes benefícios:

- **Personalização do aprendizado:** os conteúdos podem ser adaptados para atender às demandas específicas de cada turma, otimizando os processos educacionais e aprofundando a base de conhecimento dos alunos.
- **Ambiente seguro:** a criação de situações controladas, onde os dados são gerados artificialmente, garante a segurança e a privacidade dos estudantes, além de evitar qualquer discriminação. Dessa forma, não compromete a saúde dos alunos e segue princípios éticos rigorosos.
- **Uso eficiente de recursos financeiros:** a utilização de dados sintéticos elimina a necessidade de grandes centros de pesquisa, equipamentos de laboratórios caros e pessoal altamente qualificado para operá-los. Os dados gerados são tão realistas e aplicáveis quanto os dados reais, tornando o processo de ensino e aprendizagem mais acessível e econômico.
- **Democratização do acesso à informação:** os estudantes têm acesso a diversas bases de dados, abrangendo desde a geração de imagens até o modelamento de sinais para Interfaces Neurais Interativas (INI). Isso proporciona um ambiente de aprendizado enriquecedor e amplia as oportunidades de exploração e desenvolvimento de habilidades.

Desta forma, propõe-se a utilização de dados sintéticos, juntamente com uma abordagem metodológica de ensino-aprendizagem centrada no estudante, como uma das vias para a criação de um sistema educacional mais inclusivo, envolvente e eficaz, buscando nesse modelo a preparação dos alunos para os desafios e oportunidades do século XXI, impulsionando o progresso tecnológico, social e ambiental.

## REFERÊNCIAS

- ANDERSON, Theodore Wilbur; FINN, Jeremy D. **The new statistical analysis of data**. New York: Springer Science & Business Media, 2012.
- DALENIUS, Tore. Privacy transformations for statistical information systems. **Journal Of Statistical Planning And Inference**, [S.L.], v. 1, n. 1, p. 73-86, fev. 1977. Elsevier BV. [http://dx.doi.org/10.1016/0378-3758\(77\)90007-6](http://dx.doi.org/10.1016/0378-3758(77)90007-6).
- GATI, Nicholaus J. *et al.* Differentially private data fusion and deep learning Framework for Cyber-Physical-Social Systems: state-of-the-art and perspectives. **Information Fusion**, [S.L.], v. 76, p. 298-314, dez. 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.inffus.2021.04.017>.
- GUI, Jie *et al.* A Review on Generative Adversarial Networks: algorithms, theory, and applications. **Arxiv**, [S.L.], v. [], n. [], p. 1-28, jan. 2020. ArXiv. <http://dx.doi.org/10.48550/ARXIV.2001.06937>.
- MASETTO, Marcos. Inovação na Educação Superior. **Interface - Comunicação, Saúde, Educação**, [S.L.], v. 8, n. 14, p. 197-202, fev. 2004. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s1414-32832004000100018>.
- MIGUEL, Paulo Victor de Oliveira. **ECOLIG**: o protocolo semiótico para codificação e transferência do conhecimento. 2010. 1 v. Tese (Doutorado) - Curso de Engenharia Elétrica e de Computação, Unicamp, Campinas, 2009. Disponível em: <https://repositorio.unicamp.br/Busca/Download?codigoArquivo=467129>
- MORÁN, José. Mudando a educação com metodologias ativas. **Coleção Mídias Contemporâneas**. Convergências Midiáticas, Educação e Cidadania: aproximações jovens, v. 2, p. 15-33, 2015.
- MOREIRA, Daniel Augusto. Elementos para um plano de melhoria do ensino universitário a nível da instituição. **Revista Imes**, São Caetano do Sul, v. 3, n. 9, p. 28-32, 1986.
- NUÑã, Florian Marcel *et al.* The relationship among urbanization, economic growth, renewable energy consumption, and environmental degradation: a comparative view of european and asian emerging economies. **Gondwana Research**, [S.L.], v. 128, p. 325-339, abr. 2024. Elsevier BV. <http://dx.doi.org/10.1016/j.gr.2023.10.023>.
- PINHEIRO, Hildete Prisco. **O que são dados?** IMECC - UNICAMP. Disponível em: <https://www.ime.unicamp.br/~hildete/dados>. Acesso em: 10 abr. 2024.
- QUINTANILHA, Luiz Fernando. Inovação pedagógica universitária mediada pelo Facebook e YouTube: uma experiência de ensino-aprendizagem direcionado à geração-z. **Educar em Revista**, [S.L.], n. 65, p. 249-263, set. 2017. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/0104-4060.50027>.
- RAGHUNATHAN, Trivellore E.. Synthetic Data. **Annual Review Of Statistics And Its Application**, [S.L.], v. 8, n. 1, p. 129-140, 7 mar. 2021. Annual Reviews. <http://dx.doi.org/10.1146/annurev-statistics-040720-031848>.

SANTOS, Sandra Carvalho dos. O PROCESSO DE ENSINO-APRENDIZAGEM E A RELAÇÃO PROFESSOR-ALUNO: aplicação dos sete princípios para a boa prática na educação de ensino superior. **Caderno de Pesquisa em Administração**, São Paulo, v. 8, n. 1, p. 69-82, mar. 2001.

SHUJA, Junaid *et al.* COVID-19 open source data sets: a comprehensive survey. **Applied Intelligence**, [S.L.], v. 51, n. 3, p. 1296-1325, 21 set. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10489-020-01862-6>.

SKINNER, Burrhus Frederic. **The technology of teaching**. BF Skinner Foundation, 2016.

VALÉRIO, Marcelo; BAZZO, Walter Antônio. O papel da divulgação científica em nossa sociedade de risco: em prol de uma nova ordem de relações entre ciência, tecnologia e sociedade. **Revista de Ensino de Engenharia**, v. 25, n. 1, p. 31-39, dez. 2006.

VIEIRA, Ana Luzia Silva. Educação, escola e docência: novos tempos, novas atitudes, de Mario Sérgio Cortella. **Dialogia**, [S.L.], n. 20, p. 248-250, 16 dez. 2014. Universidade Nove de Julho. <http://dx.doi.org/10.5585/dialogia.n20.5355>.

VON ZUBEN, Andre; VIANA, Felipe A.C. Generative adversarial networks for extrapolation of corrosion in automobile images. **Expert Systems With Applications**, [S.L.], v. 213, p. 118849, mar. 2023. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2022.118849>.

VON ZUBEN, Fernando José. **Generative Adversarial Network (GAN)**. 2024. Disponível em: <https://www.dca.fee.unicamp.br/~vonzuben/courses/>. Acesso em: 01 maio 2024.

## INSTRUCTION TO THE STUDY OF SYNTHETIC DATA GENERATION AS A TOOL FOR THE TEACHING-LEARNING PROCESS IN ENGINEERING.

**Abstract:** *Humanity is experiencing a unique moment in its history, experiencing a gigantic range of data (Big Data), with a tendency to gather information, analyze it and reach conclusions, as well as developing algorithms and systems based on Artificial Intelligence (AI). every day more robust for the most varied applications. The growing need for Expert and Intelligent Systems demands the extraction, analysis and application of data, where concerns may sometimes arise involving privacy, integrity and lack of data. In this sense, this work illustrates that data aims to serve as a basis for students, developers and researchers to model systems and solve problems. In several situations, its extraction from the real world becomes slow and impractical, whether due to the time required for monitoring and collection, external interference or even the bias that each data brings. Therefore, with purely real data, some applications may become unfeasible. It is in this sense that the incorporation of new data generation and modeling techniques is necessary within Engineering. The generation of synthetic data is a way to guarantee the privacy, integrity and volume of information, enabling the time, quantity, integrity and privacy of the volume of data necessary for training Intelligent Systems. Furthermore, the work permeates the foundations of the relationship between the teacher and students, showing that over time, Engineering has considerably impacted the way society evolves, establishing a new technical and scientific standard. The union between science and technology in Engineering*

*is responsible for renewing hopes and fueling social expectations for the future, focusing on overcoming the most varied problems. Modern society is dependent on Engineering and technological advances are the result of this work, it is of fundamental importance to incorporate effective ways of developing skills and abilities in students, so that they can supplement social needs with a solid scientific basis, in addition to a global interpretation of more varied problems. Therefore, we propose in this work the incorporation of the generation of synthetic data through a Generative Adversarial Network (GAN), as a tool for the educational process, linked to a teaching methodological proposal, where the teacher occupies the role of guiding and the student as the protagonist. of your learning process.*

**Keywords:** *synthetic data, teaching-learning process, intelligence artificial.*

