



SÃO PAULO

FACULDADE SENAI DE TECNOLOGIA MECATRÔNICA
REVISTA BRASILEIRA DE MECATRÔNICA

MACHINE LEARNING: APLICAÇÕES NA INDÚSTRIA QUÍMICA

MACHINE LEARNING: CHEMICAL INDUSTRY APPLICATIONS

Kevy Pontes Eliodório¹

Ricardo Alexandre Carmona²

Daniel Otávio Tambasco Bruno³

RESUMO

Indústria 4.0 é o nome dado à tendência atual de automação e gerenciamento de informação nas tecnologias de manufatura, comumente referenciada como a quarta revolução industrial. Nesta nova fase, a principal mudança é na integração de tecnologias digitais e físicas, que tem potencial para alavancar o setor industrial e a competitividade das empresas frente a um mercado cada vez mais “VUCA”. Na indústria química, sabe-se que muitas reações e processos são notadamente difíceis de controlar e várias delas são afetadas por diferentes fatores operacionais. Entretanto, imagine, por exemplo, que você diga a um sistema o composto químico que deseja produzir e ele lhe dê o caminho para a síntese, todas as suas características e condições para que a reação ocorra, para isso, os principais candidatos são as técnicas de *Machine learning* (ML). ML surge com o potencial de solucionar e inovar a indústria química, reduzindo custos e ajudando, ou até mesmo, substituindo a tomada de decisão humana. Desta forma, no presente trabalho, diversas aplicações de *Machine learning* serão discutidas com o objetivo de mostrar a oportunidade que esse setor produtivo tem de alavancar sua competitividade frente a um mundo em que as alterações ocorrem a cada instante.

ABSTRACT

Industry 4.0 is the name given to the current trend of automation and information management in manufacturing technologies, commonly referred to as the fourth industrial revolution. In this new phase, the main change is in the integration of digital and physical technologies, which has the potential to improve the industrial sector and the competitiveness of companies facing an increasingly "VUCA" market. In the chemical industry is known that many reactions and processes are notably difficult to control and several of them are affected by different operational factors. However, imagine, for example, that you tell a system the chemical compound you want to produce and it will give you the path to the synthesis, all its characteristics, and conditions for the reaction to occur, in this context the main candidates are the techniques of *Machine learning* (ML). ML arises with the potential to solve and innovate the chemical industry, reducing costs and helping,

¹ Universidade de São Paulo, Escola Politécnica da USP, kevyPontes@hotmail.com

² Serviço Nacional de Aprendizagem Industrial – Departamento Regional de São Paulo, rcarmona@sp.senai.br

³ Faculdade Senai de Tecnologia Mecatrônica Industrial, daniel.bruno@sp.senai.br

or even replacing, human decision making. Thus, in the present work, several applications of *Machine learning* will be discussed with the aim of showing the opportunity that this productive sector has to leverage its competitiveness in front of a world in which the changes occur every moment.

Data de submissão: (17/02/2019)

Data de aprovação: (26/02/2019)

1 INTRODUÇÃO

1.1 Indústria 4.0 e *Machine learning*

Indústria 4.0 é o nome dado à tendência atual de automação e gerenciamento de informação nas tecnologias de manufatura, comumente referenciada como a quarta revolução industrial. Nesta nova fase, a principal mudança é na integração de tecnologias digitais e físicas, que tem potencial para alavancar o setor industrial e a competitividade das empresas frente a um mercado cada vez mais “VUCA” (Volátil, incerto, complexo e ambíguo) (WUEST et al., 2016).

A Indústria 4.0 e seus preceitos são o próximo passo no caminho para a digitalização do setor de manufatura, em que os sistemas baseados em Inteligência Artificial (IA) e computação cognitiva (CC) não estão apenas mudando a maneira como interagimos com a informação, os computadores e a produção, mas também a revolucionando e trazendo uma nova forma de entender os processos e suas interações (LEE; SHIN; REALFF, 2018).

Conceitos básicos de produção como aumento de produtividade, minimização de interferência humana e erros manuais, redução de custos produtivos, melhor aproveitamento de infraestrutura e esforço humano focado em ações não repetitivas para melhoria de eficiência são levados ao extremo nessa nova forma de produção baseada em sistemas autônomos.

A essa tendência inclui-se os sistemas cyber-físicos, a internet das coisas (*Internet of things* - IoT), computação em nuvem, mineração de dados e, por fim, computação cognitiva, que surge como um dos grandes diferenciais na produção em larga escala (figura 1).

Figura 1 - Pilares da indústria 4.0



Fonte: LWT Sistemas (2018)

Baseada principalmente na inteligência artificial e no processamento de dados, a computação cognitiva abrange áreas como aprendizado de máquinas, inferência automatizada, visão computacional e reconhecimento e processamento de fala. No contexto da Indústria 4.0, a utilização da capacidade computacional, caracterizada por capacidades não-supervisionadas de aprendizado e interação em tempo real, surge como uma ferramenta para auxiliar ou substituir a tomada de decisão humana, em que as máquinas se aproximam cada vez mais da nossa cognição.

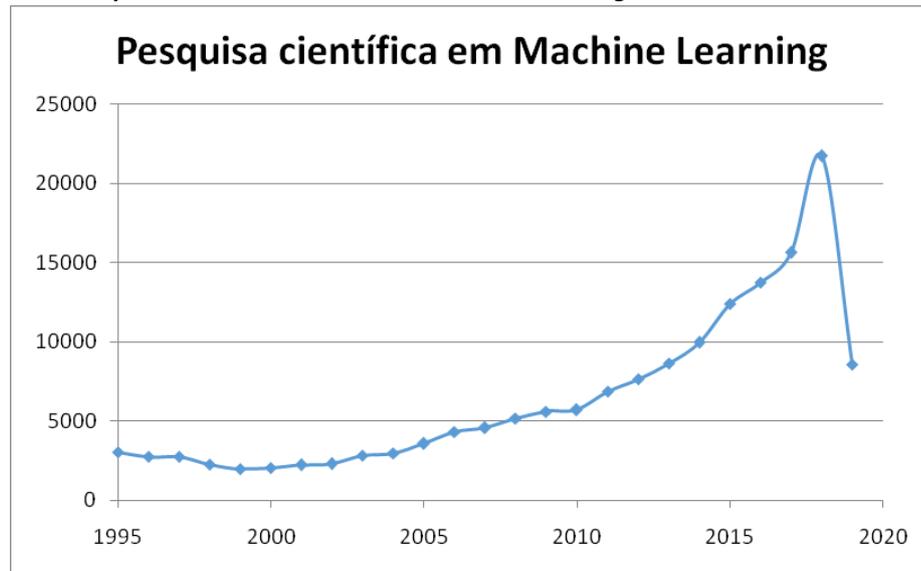
As aplicações de cognição em máquinas e sistemas podem ser as mais variadas possíveis, dentre elas destaca-se a aplicação no controle de processos. A integração do controle de processos com otimização é crítica na produção inteligente sugerida pela Indústria 4.0. Além disso, sistemas cada vez mais complexos tornam necessárias as aplicações de técnicas avançadas de controle, que ofereçam capacidade de adaptação a diferentes situações (SHAO et al., 2018).

1.2 *Machine learning* na Indústria Química

Os engenheiros químicos utilizam rotineiramente computadores para modelagem, simulação, projeto e otimização de processos químicos. Tipicamente, os métodos computacionais são algoritmos numéricos (regressão, estatística, equações diferenciais, etc.) realizados com ferramentas de software como Excel®, MatLab®, AspenPlus®, entre outros. Todos esses métodos e ferramentas são aplicações de princípios de engenharia química e seus pilares como termodinâmica, fenômenos de transporte, cinética de reações químicas e operações unitárias, que o engenheiro combina e aplica de forma ordenada para resolver problemas de alta complexidade encontrados na indústria. Mas e se os computadores pudessem fazer mais do que meros cálculos e algum dia resolver problemas como um engenheiro?

Sabe-se que muitas reações químicas e processos usualmente encontrados na indústria são notadamente difíceis de controlar devido a qualidade do produto final e várias delas são afetadas por diferentes fatores como temperatura, pH, concentração de reagentes, pureza, rota sintética, reagentes e modo de operação dos reatores. Entretanto, imagine, por exemplo, que você diga a um sistema o composto químico que deseja produzir e ele lhe dê o caminho para a síntese, todas as suas características e condições para que a reação ocorra. Seria isso possível em uma nova forma de produção baseada em *Machine learning*? (COLEY et al., 2017; SEGLER; WALLER, 2017; GAO et al., 2018).

Um forte indicativo do uso crescente de técnicas de ML como uma nova ferramenta para impulsionar a indústria é a pesquisa científica. A figura 2 mostra o crescimento na pesquisa relacionada à computação cognitiva e *Machine learning* desde as duas últimas décadas (*Science Direct*, 2019). Dados desde 1995 mostram o crescimento de em média três mil pesquisas em meados da década de 90 para mais de 20 mil pesquisas apenas no ano de 2018.

Figura 2 - Pesquisa científica relacionada a *Machine learning*

Fonte: Autor

Dentro dessas pesquisas, as aplicações são as mais diversas possíveis, entre elas, análise de capacidade de produção e otimização energética (HAN et al., 2019), controle de tráfego aéreo (VERDONK GALLEGO et al., 2018), engenharia metabólica (CAMACHO et al., 2018), Biomecânica do corpo humano (HALILAJ et al., 2018), análise psicológica de tendências suicidas (BURKE; AMMERMAN; JACOBUCCI, 2019), entre outras mais variadas possíveis.

Economicamente, a indústria química enfrenta um grande desafio: entregar lucro em um mercado global extremamente competitivo. Em vista disso, muitas indústrias fortaleceram suas posições por meio de fusões e aquisições de outras companhias. Dadas às condições instáveis na indústria química, incerteza é a nova condição normal.

A indústria química está antecipando oportunidades ao agregar a tecnologia de aprendizado de máquinas (ML). A utilização do ML pode oferecer oportunidades na indústria química para fazer mais com menos pessoas, mais segurança e reduzir o custo de administrar o negócio, gerando um salto produtivo necessário e adequado a produção na era 4.0 (SADIKU et al., 2017).

2 REVISÃO BIBLIOGRÁFICA

2.1 Desafios do setor manufatureiro

A indústria de manufatura é uma indústria muito estabelecida e robusta, no entanto, até hoje a sua importância não pode ser quantificada de maneira adequada. Várias economias mundiais experimentaram uma redução da manufatura na contribuição para o produto interno bruto (PIB) nas últimas décadas, resultado de crises, um mercado cada vez mais instável e rápidas mudanças no setor (WUEST et al., 2016).

Nos últimos anos, contudo, várias iniciativas de impulsionamento do setor manufatureiro se iniciaram. Exemplos são os EUA através de ações como '*Executive Actions to Strengthen Advanced Manufacturing in America*' (Ações Executivas para Fortalecer a Manufatura Avançada na América) e a União Europeia com as suas "Fábricas do Futuro". Os

desafios que a fabricação enfrenta hoje são diferentes dos desafios encontrados no passado. A seguir destacam-se os principais (SHIANG; NAGARAJ, 2011; THOMAS; BYARD; EVANS, 2012; WUEST et al., 2016):

- a) adoção de tecnologias avançadas de fabricação;
- b) importância crescente da fabricação de produtos de alto valor agregado;
- c) utilização de conhecimento avançado, gerenciamento de informações e sistemas de inteligência artificial;
- d) processos e produtos de manufatura sustentável;
- e) recursos empresariais e cadeias de suprimentos ágeis e flexíveis;
- f) inovação em produtos, serviços e processos;
- g) colaboração estreita entre indústria e pesquisa para adotar novas tecnologias;
- h) novos paradigmas de gestão de manufatura.

Esses desafios destacam a tendência contínua de mudança da manufatura para um sistema produtivo mais complexo e dinâmico. A aparente complexidade é herdada não só dos programas de fabricação em si, mas cada vez mais do produto a ser fabricado, fruto das exigências do mercado, além das relações de negócio das empresas e redes de colaboração. Somando-se a esse desafio está o fato de que o ambiente dinâmico de negócios das empresas de manufatura de hoje é fortemente afetado pela incerteza econômica.

Olhando especialmente para domínios com maior probabilidade de serem otimizados, por exemplo, monitoramento, controle, programação e diagnósticos, torna-se evidente que a crescente disponibilidade de dados está adicionando outros desafios: as próprias grandes quantidades de dados e informações disponíveis (por exemplo, dados de sensor), a alta dimensionalidade e variedade de dados (por exemplo, devido a diferentes sensores ou processos conectados), bem como a natureza complexa das relações estabelecidas nos processos e problemas de otimização de fabricação que tornam cada vez mais difícil a análise e controle da produção. Ao mesmo tempo em que adiciona desafios, a alta disponibilidade de dados abre espaço para oportunidades jamais vistas até então, incluindo o distanciamento humano das atividades (WUEST; IRGENS; THOBEN, 2014).

Para superar alguns dos principais desafios atuais dos sistemas de manufatura complexos, os candidatos são técnicas de aprendizado de máquina. Essas abordagens baseadas em dados são capazes de encontrar padrões altamente complexos e não lineares em dados de diferentes tipos e fontes e transformar dados brutos para espaços de características, os chamados modelos, que são então aplicados para previsão, detecção, classificação, regressão ou previsão.

2.1.1 Aprendizado de máquina frente aos atuais desafios de fabricação

Antes de examinar o aprendizado de máquina (ML) com base nos resultados requeridos para uma solução do futuro da manufatura, algumas definições usadas são brevemente apresentadas. O *Machine learning* é conhecido por sua capacidade de lidar com problemas de natureza complexa, que muitas vezes aparecem nos sistemas de manufatura inteligente (WUEST et al., 2016).

A aplicação de técnicas de ML aumentou nas últimas duas décadas devido a vários fatores, por exemplo, a disponibilidade de grandes quantidades de dados complexos, maior capacidade de processamento dos sistemas, além da maior capacidade de utilização e poder das ferramentas de ML disponíveis (SMOLA; VISHWANATHAN, 2008).

Atualmente, o ML é amplamente aplicado em diferentes áreas da fabricação, por exemplo, otimização, controle, análise de dados e solução de problemas diversos (ALPAYDIN, 2010).

Muitas técnicas ML são projetadas para analisar grandes quantidades de dados e são capazes de lidar muito bem com alta variedade das informações e alta dimensionalidade (termo que se refere ao fenômeno que surge durante a análise e organização de dados em espaços de alta dimensões). No entanto, questões relacionadas, como possíveis sobreajustes, devem ser consideradas durante a aplicação (ALPAYDIN, 2010).

A aplicação de ML na manufatura pode resultar em um conjunto de padrões a partir de dados existentes, que podem fornecer uma base para o desenvolvimento de relações e aproximações sobre o comportamento futuro do sistema. Esta nova informação pode auxiliar na tomada de decisão humana sobre o processo (supervisionado) ou ser usado automaticamente para melhorar diretamente o sistema (não supervisionado). No final, o objetivo das técnicas de ML é detectar certos padrões ou regularidades que descrevem as relações complexas existentes (ALPAYDIN, 2010; WUEST et al., 2016).

Dado o desafio de um ambiente de produção com rápidas mudanças, o ML herda a capacidade de aprender e se adaptar às mudanças e, assim, deve fornecer soluções para todas as situações possíveis. Torna-se claro, desta maneira, por que sua aplicação na fabricação pode ser benéfica dada a dificuldade da maioria dos modelos mais comuns para lidar com a adaptabilidade (WUEST et al., 2016).

No entanto, uma análise mais detalhada das técnicas de ML disponíveis, bem como os seus pontos fortes e limitações relativas aos requisitos da indústria devem ser fornecidas. Há muitas perguntas a serem respondidas, como, por exemplo, as técnicas de ML podem resultar em informação qualitativa em meio a tanta informação disponível, relevante ou não.

2.2 Algoritmo e técnicas de aprendizado de máquinas

O aprendizado de máquina tornou-se um campo amplo e diversificado de pesquisa, isso levou a uma variedade de diferentes subdomínios, algoritmos, teorias e áreas de aplicação. A relação e estrutura entre os diferentes elementos não são comumente acordados e diferentes pesquisadores escolhem abordagens variadas para estruturar a área.

Essa estrutura destaca a importância da diferenciação de tarefa (qual é o objetivo) e algoritmo (como esse objetivo pode ser alcançado) dentro do campo ML. Dentro dos algoritmos, três classes são encontradas:

- a) aprendizado por reforço: menos *feedback* é dado, mas apenas uma avaliação da ação escolhida é dada pelo professor;
- b) aprendizado não supervisionado: nenhuma avaliação da ação é fornecida, uma vez que não há professor;
- c) aprendizado supervisionado: a resposta correta é fornecida por um professor.

Essa estrutura é amplamente aceita, no entanto, ainda existem divergências em relação ao que se encaixa nessas três classes. Por exemplo, alguns autores mapeiam supervisionado, não supervisionado e RL como uma subdivisão de Redes Neurais (NN).

Os diferentes algoritmos e abordagens combinatórias frequentemente tendem a ser adaptados para problemas específicos, dificultando sua comparação. No entanto, uma abordagem mais promissora para selecionar um algoritmo adequado é procurar problemas de natureza semelhante e analisar qual algoritmo ML foi usado para resolvê-lo e quais os

resultados. Este é um bom ponto de partida. Depois que o algoritmo é aplicado o problema e os primeiros resultados estão disponíveis, diferentes métodos podem ser aplicados e os resultados para o problema dado podem ser comparados (WUEST et al., 2016).

A seguir, aprendizado de máquina não supervisionado, RL e aprendizado de máquina supervisionado são brevemente descritos para permitir diferenciá-los uns dos outros.

2.2.1 Aprendizado de máquina não supervisionado

O aprendizado de máquina não supervisionado é outra grande área de pesquisa. O atributo definidor é que dentro de aprendizagem não supervisionada, não há feedback de um chamado “professor externo”, o próprio algoritmo deve identificar e agir a partir de dados existentes. O objetivo é descobrir classes desconhecidas de itens por agrupamento, enquanto a aprendizagem supervisionada é focada na classificação (rótulos previamente conhecidos) (ALPAYDIN, 2010).

Basicamente, o ML não supervisionado descreve qualquer processo que tenta aprender uma estrutura na ausência de uma saída identificada (ML supervisionado) ou *feedback* (RL). Três exemplos típicos de aprendizado não supervisionado são agrupamentos (*clustering*), regras de associação e mapas auto-organizáveis. (ALPAYDIN, 2010).

Especialmente no contexto de Big Data, os métodos não supervisionados estão se tornando cada vez mais importantes e surgem como uma solução promissora. No entanto, como na aplicação de manufatura, a principal suposição é que o conhecimento especialista ou o professor podem fornecer *feedback* sobre a classificação dos estados para identificar a aprendizagem definida em vez de treinar o algoritmo e, assim, o foco será realocado em métodos supervisionados (WUEST et al., 2016).

Alguns aspectos da aprendizagem não supervisionada, entretanto, podem ser benéficos na aplicação de fabricação. Em primeiro lugar, existe a possibilidade de que, em alguns casos, pode não haver esse *feedback* especializado disponível. Outro aspecto é realizar abordagens híbridas aplicando um pouco de cada metodologia, abordagem que ganha importância devido ao rápido aumento de dados não rotulados em ambientes fabris (KANG; KIM; CHO, 2016). E por último, métodos não supervisionados podem e estão sendo usados para, por exemplo, identificar *outliers* (dados fora do padrão) em dados de fabricação (HANSSON et al., 2016).

2.2.2 Aprendizado de reforço

Aprendizado de reforço ou *Reinforcement Learning* (RL) é definido pelo aprendizado cuja informação de treinamento é fornecida pelo ambiente. A informação de quão boa foi a performance do sistema é fornecida por um sinal de reforço (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2007). Outra característica que define é que o sistema tem que descobrir tentando quais ações geram os melhores resultados (sinal de reforço numérico), em vez de ser dito, característica que distingue RL de outras metodologias. No entanto, mesmo com essas características, alguns pesquisadores entendem a RL como uma categoria do aprendizado supervisionado, enquanto outros caracterizam como uma subdivisão de Redes Neurais (WUEST et al., 2016).

No entanto, diferente da aprendizagem supervisionada problemas de RL podem ser descritos pela ausência de exemplos conhecidos de comportamentos, esperados ou não (ALPAYDIN, 2010). RL, com base na resposta sequencial do ambiente, emula o processo de

aprendizagem de seres humanos. Este "sinal de recompensa", que pode ser percebido na RL diferencia de algoritmos não supervisionados. Diferente da aprendizagem supervisionada, a RL é mais adequada em situações em que não há "supervisor", fazendo necessário o uso de um agente que possa aprender com a interação e sua própria experiência. Como a RL é baseada no *feedback* das ações, uma questão interessante e também desafiadora é que certas ações não têm apenas um impacto imediato, mas certos efeitos podem aparecer mais tarde ou durante uma avaliação adicional seguinte.

No geral entende-se que a RL é definida não caracterizando metodologias de aprendizagem, mas caracterizando um problema de aprendizagem e, assim, qualquer método adequado para resolver esse problema, pode ser considerado um método de reforço de aprendizagem (ALPAYDIN, 2010).

Um desafio muito específico para a RL é a troca entre o que foi explorado e o que será explorado. Para atingir o objetivo, o agente tem que explorar as ações que aprendeu para identificar aqueles que têm que explorar, tentando ativamente novas maneiras.

2.2.3 Aprendizado supervisionado

Na aplicação de manufatura, as técnicas supervisionadas de ML são as mais aplicadas principalmente devido à riqueza de dados e conhecimento escasso sobre a natureza dos problemas. Além disso, ML supervisionado pode se beneficiar da coleta de dados durante a fabricação em processos de controle.

Basicamente, o ML supervisionado está aprendendo com exemplos fornecidos por um supervisor. Isso se deve em parte à disponibilidade de *feedback* especializado e dados rotulados. O ML supervisionado é aplicado em diferentes domínios da fabricação, monitoramento e controle, sendo muito promissor nestas áreas (ALPAYDIN, 2010; KWAK; KIM, 2012; WUEST; IRGENS; THOBEN, 2014).

O processo geral do ML supervisionado contém várias etapas para lidar com os dados, o conjunto de dados de treinamento e teste pelo professor. Baseada em um determinado problema, os dados necessários são identificados e se necessário pré-processados (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2007).

Um aspecto importante é a definição do conjunto de treinamento, pois este influencia fortemente os resultados. Mesmo assim, muitas vezes parece que a seleção do algoritmo está sempre seguindo a definição do conjunto de dados de treinamento, a definição dos dados de treinamento também deve considerar os requisitos do algoritmo. Alguns algoritmos permitem a chamada '*kernel selection*' para adaptar o algoritmo à natureza específica do problema, aqui se destaca a adaptabilidade do ML e a variedade de problemas que podem ser resolvidos (WUEST et al., 2016).

Requisitos semelhantes são até certo ponto verdadeiros para a identificação e pré-processamento dos dados uma vez que algoritmos diferentes têm certas vantagens e desvantagens relativas à manipulação de diferentes conjuntos de dados (por exemplo, formato, dimensões, etc.).

Depois que um algoritmo é selecionado, ele é treinado usando o conjunto de dados de treinamento. A fim de julgar a capacidade de realizar a tarefa alvo, o algoritmo treinado é então avaliado usando um conjunto de dados de avaliações ou teste. Dependendo do desempenho do algoritmo treinado com o conjunto de dados de avaliação, os parâmetros podem ser ajustados para otimizar o desempenho, caso o desempenho já seja bom. Caso o desempenho não seja satisfatório, o processo deve ser reiniciado em um estágio anterior,

dependendo do desempenho real obtido. A regra convencional é de 70% dos dados usados como treinamento, 20% avaliação e ajuste de parâmetros (*bias*, por exemplo) e 10% como teste (ALPAYDIN, 2010).

Dentre as metodologias existentes dos algoritmos supervisionados, as redes neurais artificiais (Artificial Neural Networks- ANNs) são as mais aplicadas, além delas, destacam-se:

- a) Statistical Learning Theory (STL);
- b) Redes bayesianas (BNs);
- c) Memory-Based Reasoning (MBR);
- d) Instance-Based Learning (IBL);
- e) Lógica Fuzzy;
- f) Algoritmos genéticos;

Entre outras técnicas existentes, as características de cada uma não serão tratadas aqui, entretanto, diversos trabalhos reportam suas vantagens e aplicações com mais detalhes.

3 ESTADO DA ARTE

Neste capítulo, aplicações serão apresentadas tentando apresentar ML como uma solução interessante e promissora na indústria química. As aplicações reportadas nesta seção não são uma tentativa de generalizar, ao contrário, o objetivo é apresentar e exemplificar possíveis soluções para a indústria química baseadas em *Machine learning*.

3.1 Predição de reações orgânicas aplicando *Machine learning*

Um trabalho recente foi selecionado para estudo de caso da aplicação de técnicas de ML na tentativa de predição de reações químicas. O estudo descreve um modelo, baseado em reações rígidas e *Machine learning*, que aprende a prever os principais produtos de reações orgânicas, dado um conjunto de moléculas reagentes (COLEY et al., 2017).

O trabalho contribui apresentando uma estratégia de aumento de dados através dos quais bancos de dados são suplementados com exemplos de reações quimicamente plausíveis disponíveis na literatura. O resultado dessa estratégia aplicada automaticamente, no qual poucas informações estão disponíveis e não é necessária interferência humana, é apresentado.

O estudo ainda contribui com uma nova representação de reação focada nas transformações baseadas em pontos específicos (sítios) dos compostos químicos, em vez de simplesmente reagentes e produtos. Os autores apresentam os resultados da implementação e validação de um modelo baseado em redes neurais que aprende quando certos modos de reatividade são mais propensos de ocorrer que outros, inclusive com reações de baixo rendimento.

A assistência computacional nos problemas de síntese é utilizada há mais de 40 anos, entretanto, a aplicação de softwares de planejamento ainda enfrenta diversos problemas para uma adoção generalizada. Um dos maiores desafios no desenvolvimento de uma rota é que na maioria das vezes ela falha no laboratório, apesar de inicialmente parecer viável. A verdadeira resposta de sucesso de um programa de síntese é se ele consegue modelar e prever o que é obtido laboratorialmente.

Assim, usando 15.000 registros de reações experimentais de patentes concedidas dos Estados Unidos, os autores apresentam um modelo que foi treinado para selecionar os

principais produtos registrados, classificando uma lista de candidatos gerada pelo próprio modelo, em que um é conhecido como o principal produto. Reações candidatas são representadas usando transformações fundamentais dos reagentes para produtos nos seus sítios de reação, em vez das estruturas globais convencionais das moléculas constituintes.

Os resultados mostram que em uma validação cruzada de 5 tentativas, o modelo treinado atribui o produto principal chamado rank 1 em 71,8% dos casos, rank ≤ 3 em 86,7% dos casos e rank ≤ 5 em 90,8% dos casos, sendo bastante representativo como estratégia de base para a predição de produtos de reação.

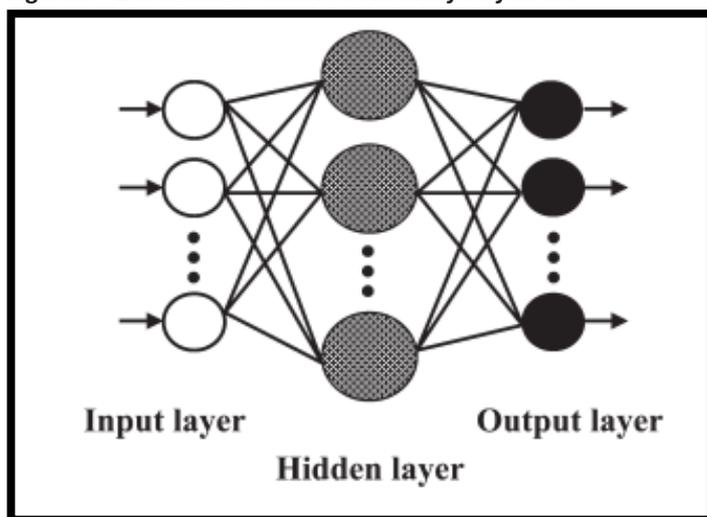
Assim, existe uma grande possibilidade do aprendizado de máquina se tornar protagonista no design de síntese assistido, não só como um assistente automatizado de planejamento, mas como uma ferramenta independente para químicos e engenheiros avaliarem a viabilidade da reação. O trabalho utilizado como estudo de caso representa um passo significativo no futuro da química pela aplicação de soluções de ML, como uma ferramenta para triagem de reações e validações, complementando a química orgânica experimental.

3.2 Uso de redes neurais para predição de condutividade térmica de soluções eletrolíticas

Outro trabalho interessante utilizado como estudo de caso é o de Eslamloueyan, Khademi e Mazinani (2011). Neste estudo, uma rede perceptron multicamada (*Multi layer Perceptron - MLP*) é proposta para prever a condutividade térmica de uma solução eletrolítica.

A MLP é uma classe das redes neurais *feedforward*, em que sua estrutura se baseia em três camadas distintas: uma camada de entrada (inputs), uma camada escondida e uma camada de saída (outputs). Cada uma delas, a exceção da camada de entrada, usa uma função de ativação não linear (figura 3).

Figura 3 - Estrutura de uma rede neural *feedforward*



Fonte :Eslamloueyan, Khademi e Mazinani (2011).

Os autores avaliaram a predição do modelo em uma ampla faixa de temperaturas e concentrações baseadas no peso molecular e número de elétrons do soluto, mostrando ainda mais como as redes neurais (*Machine learning*) são capazes de lidar com altas dimensões e dados que se relacionam de maneira complexa.

A precisão da rede neural artificial proposta foi avaliada através da análise de regressão dos valores previstos e experimentais de várias soluções aquosas, incluindo algumas das quais não foram utilizadas no treinamento da rede.

Os resultados apresentados e a comparação da rede MLP desenvolvida com outras correlações recomendadas na literatura indicam que a arquitetura da rede neural proposta supera outros métodos alternativos, tanto em termos de precisão quanto de extrapolação e predição. Além disso, as correlações de condutividade de outros geralmente são sugeridas para soluções eletrolíticas específicas e uma faixa limitada de temperaturas e concentrações, enquanto tais limitações não existem para a rede MLP proposta.

Desta forma, novamente, as técnicas de *Machine learning* se apresentam como uma solução aplicável que pode lidar com os principais problemas industriais e técnicos encontrados durante a modelagem, controle e operação.

3.3 Controle de processos químicos por redes neurais

Para exemplificar aplicações de ML no controle de processos, foi selecionado o trabalho de Ahmed (2013). O trabalho analisado como estudo de caso reporta um esquema de controle baseado em redes neurais artificiais para o estudo do controle de reator contínuo de tanque agitado (CSTR), coluna de destilação e processo de neutralização e este método é comparado com o controlador proporcional, integral e derivativo (PID) convencional.

Do ponto de vista industrial, as operações unitárias selecionadas são algumas das mais comuns no dia-a-dia dos engenheiros químicos, além disso, o autor compara as soluções baseadas em computação cognitiva com os controles amplamente aplicados na indústria (PID).

Um modelo de controle multicamada com propagação para trás (*multi-layer back-propagation*) foi utilizado para modelar as relações extremamente não-lineares entre entradas e variáveis controladas encontradas neste tipo de operação industrial. Além disso, o modelo permite regular as variáveis em diferentes condições de operação com uma habilidade de aprender mais flexível. O formalismo matemático envolvido no modelo, que não será abordado no presente estudo de caso, é apresentado pelo autor.

O autor estudou a robustez da técnica alterando condições do *set-point* (controle servo) e aplicando distúrbios (controle regulatório), ambos encontrados na prática industrial. Os equipamentos utilizados nas simulações dos processos industriais são montados laboratorialmente e têm suas principais variáveis monitoradas (por exemplo, temperatura do reboiler, dos pratos e do topo da coluna de destilação).

Os resultados apresentados sugerem que os controladores baseados na abordagem de redes neurais apresentam bom desempenho no rastreamento de *set-point* e na eliminação de distúrbios, com maior velocidade de resposta e um offset (desvio em relação ao *set-point* depois de atingido o estado estacionário), em média, menor que os controladores convencionais. Outras características como menor oscilação e menor tempo de estabilização das variáveis controladas também foram observadas nos experimentos para os três processos estudados.

Assim, em conclusão, as técnicas de *Machine learning control* podem apresentar uma solução eficiente para os problemas encontrados na indústria em relação aos controladores convencionais. Entretanto, a técnica utilizada na estratégia de controle não pode ser

generalizada e outros critérios diversos, como custo de implementação e complexidade exigida, devem ser ponderados na escolha do controlador ideal (AHMED, 2013).

4 CONCLUSÃO

As principais características do aprendizado de máquina foram apresentadas, seus desafios e vantagens, bem como a maneira com que essas técnicas podem se inserir em um contexto cada vez mais propenso a rápidas mudanças em uma nova forma de produção 4.0.

O presente trabalho mostrou também, por meio de estudos específicos, que as técnicas de aprendizado de máquina são ferramentas promissoras para resolver problemas na indústria química, auxiliando na tomada de decisão e no controle de processos. O aprendizado de máquina também é mostrado como uma alternativa promissora ao que se encontra disseminado na indústria e uma forma de tornar a produção cada vez mais independente da ação humana.

Além disso, o trabalho reforça cada vez mais a necessidade de um caráter interdisciplinar na indústria, unindo áreas como automação e controle a outras áreas distintas, como citado no desenvolvimento deste trabalho.

REFERÊNCIAS

AHMED, D. Artificial Neural Network Control of Chemical Processes. **Eng. & Tech. Journal**, v. 32, n. 1, p. 176–196, 2013.

ALPAYDIN, E. **Introduction to Machine Learning**, 2^o Ed., Cambridge, 2010.

BURKE, T. A.; AMMERMAN, B. A.; JACOBUCCI, R. The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. **Journal of Affective Disorders**, v. 245, p. 869–884, 2019. DOI:

<https://doi.org/10.1016/j.jad.2018.11.073>. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0165032718317506?via%3Dihub>.

Acesso em: 18 Fev. 2019.

CAMACHO, D. M. et al. Next-Generation Machine Learning for Biological Networks. **Cell**, v. 173, n. 7, p. 1581–1592, 2018. DOI: <https://doi.org/10.1016/j.cell.2018.05.015>. Disponível em: [https://www.cell.com/cell/fulltext/S0092-8674\(18\)30592-](https://www.cell.com/cell/fulltext/S0092-8674(18)30592-0?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867418305920%3Fshowall%3Dtrue)

[0?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867418305920%3Fshowall%3Dtrue](https://www.cell.com/cell/fulltext/S0092-8674(18)30592-0?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0092867418305920%3Fshowall%3Dtrue). Acesso em: 18 Jan. 2019.

COLEY, C. W. et al. Prediction of Organic Reaction Outcomes Using Machine Learning. **ACS Central Science**, v. 3, n. 5, p. 434–443, 2017.

ESLAMLOUEYAN, R.; KHADEMI, M. H.; MAZINANI, S. Using a Multilayer Perceptron Network for Thermal Conductivity Prediction of Aqueous Electrolyte Solutions. **Industrial and Engineering Chemistry Research**, v. 50, p. 4050–4056, 2011.

GAO, H. et al. Using Machine Learning to Predict Suitable Conditions for Organic Reactions. **ACS Central Science**, v. 4, n. 11, p. 1465–1476, 2018.

HALILAJ, E. et al. Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. **Journal of Biomechanics**, v. 81, p. 1–11, 2018. DOI: <https://doi.org/10.1016/j.jbiomech.2018.09.009>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0021929018307309?via%3Dihub>. Acesso em: 22 Jan. 2019.

HAN, Y. et al. Production capacity analysis and energy optimization of complex petrochemical industries using novel extreme learning machine integrating affinity propagation. **Energy Conversion and Management**, v. 180, p. 240–249, 2019. DOI: <https://doi.org/10.1016/j.enconman.2018.11.001>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0196890418312378?via%3Dihub>. Acesso em: 18 Jan. 2019.

HANSSON, K. et al. Machine Learning Algorithms in Heavy Process Manufacturing. **American Journal of Intelligent Systems**, v. 6, n. 1, p. 1–13, 2016.

KANG, P.; KIM, D.; CHO, S. Semi-supervised support vector regression based on self-training with label uncertainty: An application to virtual metrology in semiconductor manufacturing. **Expert Systems with Applications**, v. 51, p. 85–106, 2016. DOI: <http://dx.doi.org/10.1016/j.eswa.2015.12.027>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417415008295?via%3Dihub>. Acesso em: 18 Jan. 2019.

KOTSIANTIS, S. B.; ZAHARAKIS, I. D.; PINTELAS, P. E.. Supervised machine learning: A review of classification techniques. **Informatica**, v. 31, p. 249–268, 2007.

KWAK, D. S.; KIM, K. J. A data mining approach considering missing values for the optimization of semiconductor-manufacturing processes. **Expert Systems with Applications**, v. 39, n. 3, p. 2590–2596, 2012. DOI: <http://dx.doi.org/10.1016/j.eswa.2011.08.114>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417411012450?via%3Dihub>. Acesso em: 18 Jan. 2019.

LEE, J. H.; SHIN, J.; REALFF, M. J. Machine learning: Overview of the recent progresses and implications for the process systems engineering field. **Computers and Chemical Engineering**, v. 114, p. 111–121, 2018. DOI: <http://dx.doi.org/10.1016/j.compchemeng.2017.10.008>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0098135417303538?via%3Dihub>. Acesso em: 18 Jan. 2019.

LWT SISTEMAS. **Conheça os 10 pilares da indústria 4.0**. 2018. Disponível em: <https://www.lwtsistemas.com.br/10-pilares-da-industria-4-0/>. Acesso em: 18 Jan. 2019.

SADIKU, M. N. O. et al. Machine Learning in Chemical Industry. **International Journal of Advance in Scientific Research and Engineering**, p. 12–15, 2017.

SEGLER, M. H. S.; WALLER, M. P. Neural-Symbolic Machine Learning for Retrosynthesis and Reaction Prediction. **Chemistry**, v. 23, n. 25, p. 5966–5971, 2017.

SHAO, G. et al. Standards-based integration of advanced process control and optimization. **Journal of Industrial Information Integration**, p. 1–12, 2018. DOI: <https://doi.org/10.1016/j.jii.2018.10.001>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2452414X18301316?via%3Dihub>. Acesso em: 18 Jan. 2019.

SHIANG, L. E.; NAGARAJ, S. Impediments to innovation: Evidence from Malaysian manufacturing firms. **Asia Pacific Business Review**, v. 17, n. 2, p. 209–223, 2011.

SMOLA, A.; VISHWANATHAN, S. V. N. Introduction to Machine Learning. Cambridge, 2008.

THOMAS, A. J.; BYARD, P.; EVANS, R. Identifying the UK's manufacturing challenges as a benchmark for future growth. **Journal of Manufacturing Technology Management**, v. 23, n. 2, p. 142–156, 2012.

VERDONK GALLEGO, C. E. et al. Analysis of air traffic control operational impact on aircraft vertical profiles supported by machine learning. **Transportation Research Part C: Emerging Technologies**, v. 95, n. March 2018, p. 883–903, 2018. DOI: <https://doi.org/10.1016/j.trc.2018.03.017>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0968090X18303747?via%3Dihub>. Acesso em: 18 Jan. 2019.

WUEST, T. *et al.* Machine learning in manufacturing: Advantages, challenges, and applications. **Production and Manufacturing Research**, v. 4, n. 1, p. 23–45, 2016. DOI: <http://dx.doi.org/10.1080/21693277.2016.1192517>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/21693277.2016.1192517>. Acesso em: 18 Jan. 2019.

WUEST, T.; IRGENS, C.; THOBEN, K. D. An approach to monitoring quality in manufacturing using supervised machine learning on product state data. **Journal of Intelligent Manufacturing**, v. 25, n. 5, p. 1167–1180, 2014.

AGRADECIMENTOS

Os autores gostariam de agradecer a Marina Saviolli Reis da Silva e Andréia de Araújo Morandim-Glanetti pelo auxílio com a revisão técnica, bem como a Paula Saviolli Nogueira pelo auxílio na revisão ortográfica deste artigo.

SOBRE OS AUTORES**KEYVY PONTES ELIODÓRIO**

Possui graduação em Engenharia Química pela Fundação Educacional Inaciana Padre Sabóia de Medeiros (2015). Possui mestrado em Engenharia Química pela Fundação Educacional Inaciana Padre Sabóia de Medeiros (2018). Tem experiência na área de Engenharia Química, com ênfase em Processos Industriais de Engenharia Química. Atualmente é doutorando em Engenharia Química pela Escola Politécnica da Universidade de São Paulo, atuando na área de reações químicas.

CV: <http://lattes.cnpq.br/1201240377912504>

RICARDO ALEXANDRE CARMONA

Mestre em Engenharia Biomédica pela Universidade de Mogi das Cruzes, desenvolveu trabalho na subárea de Reabilitação com o título Andador Universal com Sistemas de Apoios e Controle da Aceleração. Pós-graduado em Planejamento, Implementação e Gestão da EaD pela Universidade Federal Fluminense, graduado em Engenharia Mecatrônica pela Universidade de Mogi das Cruzes, Técnico em Eletrotécnica, pela Escola Técnica Estadual Presidente Vargas de Mogi das Cruzes. Atualmente atua como Coordenador do Curso de Pós-graduação em Instalações Elétricas em B.T. e M.T., na Universidade Estácio de Sá do Rio de Janeiro, Coordenador do Curso de Pós-graduação em Automação Industrial na USCS, bem

como Especialista em Educação Profissional no SENAI-SP na gestão das 34 Unidades SENAI da Capital e Grande São Paulo, também atua como professor nos Cursos de Engenharia da FMU e Centro Universitário Estácio Radial de São Paulo, Atua também como Consultor Educacional, Parecerista Técnico, Revisor e Elaborador de Material Didático Impresso - MDI, bem como Elaborador de Itens para as Empresas Geekie e Avalia Educacional, nas áreas de Matemática, Informática, Eletrotécnica e Eletroeletrônica. Atuando mais de dois anos como tutor de cursos de pós-graduação na área de Tecnologias Digitais da Informação e Comunicação.

CV: <http://lattes.cnpq.br/0705523553879247>

DANIEL OTÁVIO TAMBASCO BRUNO

Doutorando e Mestre em Engenharia da Informação pela Universidade Federal do ABC (2013). Especialista em Análise, desenvolvimento de Sistemas e Banco de Dados pela Universidade de Ribeirão Preto (2007), Especialista em Educação a Distância pela Universidade Paulista (2012). Bacharel em Análise de Sistemas pela Universidade Paulista (2003). Atualmente é Técnico em Manufatura Digital e Professor da Faculdade SENAI de Tecnologia Mecatrônica Industrial. Tem experiência na área Gestão de Tecnologia da Informação, desenvolvimento de sistemas de informação e ciência da computação, com ênfase em processamento de imagens, inteligência artificial e Internet das coisas.

CV: <http://lattes.cnpq.br/3491851270517427>