

Uma Revisão de Literatura sobre o Uso de Aprendizado de Máquina para Distribuição de Fármacos

Helder Pestana¹, Rodrigo Bonacin^{1,2}, Ferruccio de Franco Rosa^{1,2}, Mariângela Dametto²

¹ UNIFACCAMP
Campo Limpo Paulista– SP– Brasil

² CTI Renato Archer
Campinas– SP – Brasil

h.pestana@hotmail.com, rodrigo.bonacin@faccamp.br,
ferruccio.rosa@faccamp.br, mdametto@cti.gov.br

Abstract. *The pharmaceutical formulation in the traditional way often relies on costly and uncertain trial-and-error processes. Incorporating machine learning techniques may reduce both the time and financial resources invested in pharmaceutical product development. By using machine learning techniques, computers can autonomously learn from extensive pharmaceutical, chemical, and biomedical datasets to make informed suggestions. We present a literature review on the use of machine learning in the realm of pharmaceutical formulation creation and selection with a focus on drug delivery.*

Resumo. *A formulação de medicamentos da maneira tradicional geralmente depende de processos caros e incertos, baseados em tentativa e erro. A incorporação de técnicas de aprendizado de máquina pode reduzir o tempo e os recursos financeiros investidos no desenvolvimento de produtos farmacêuticos. Usando técnicas de aprendizado de máquina, os computadores podem aprender autonomamente com extensos conjuntos de dados farmacêuticos, químicos e biomédicos para fazer sugestões informadas. Este artigo apresenta uma revisão de literatura sobre o uso de aprendizado de máquina no contexto da criação e seleção de formulações farmacêuticas com foco em distribuição de fármacos.*

Palavras-chave: Aprendizado de Máquina, Distribuição de Fármacos, Formulações Farmacêuticas, Bases de Dados Biomédicos

1. Introdução

As formulações farmacêuticas envolvem um conjunto grande de fatores, como estabilidade e biodisponibilidade, e a abordagem tradicional para o desenvolvimento destas formulações depende de processos incertos baseados em tentativa e erro, exigindo assim um grande número de experimentos *in vitro* e *in vivo*, que consomem muitos recursos e tempo [Bannigan et al. 2021]. Um outro fator preponderante na indústria farmacêutica é o rápido crescimento dos avanços tecnológicos e da quantidade de dados científicos, o que cria vários desafios tais como armazenamento e análise de dados [Kamerzell e Middaugh 2020].

Neste cenário, o aprendizado de máquina (*ML – Machine Learning*) tem se destacado como uma ferramenta para otimizar e acelerar o desenvolvimento de formulações farmacêuticas e aproveitar todo esse volume de dados para identificar padrões e tomar decisões. Entre as propriedades farmacêuticas está a distribuição de fármacos, ou seja, capacidade de transportar um fármaco (substância responsável pelo efeito terapêutico) para o seu alvo e assim alcançar um efeito terapêutico desejado.

Ao executar a revisão foram encontrados 2 trabalhos relacionados, ou seja, duas revisões sobre o tema. Bannigan et. al. (2021) apresentam uma análise sobre os métodos e ferramentas aplicadas ao desenvolvimento de formulações químicas utilizando ML. No entanto, a revisão não aborda foca no tema de distribuição de fármacos. Já Wang et al. (2021), em sua revisão, propõem o uso de IA e algoritmos de ML juntamente com modelagem molecular, modelagem matemática, simulação de processo e modelagem farmacocinética com base fisiológica para o desenvolvimento da técnica de distribuição de fármacos.

Em função de aplicações e o potencial do uso de ML em distribuição de fármacos, neste resumo estendido é apresentada uma revisão preliminar com o objetivo de identificar as técnicas e teorias utilizadas, os aspectos positivos e limitações dos estudos existentes.

2. Metodologia e Parâmetros da Revisão

Esta revisão de literatura é baseada no *guideline* apresentado por Kitchenham (2004), tendo como questão principal (QP) de pesquisa: “Quais abordagens baseadas em ML estão voltadas a otimizar distribuição de fármacos?” Os parâmetros da pesquisa foram definidos após uma análise exploratória preliminar, baseada na QP. Definiu-se o período de busca de 2019 a 2023 por ser um tema relativamente recente. A seguinte *string* de busca foi utilizada (adaptada a sintaxe de cada base): “(“*formulation*”) AND (“*machine learning*”) AND (“*pharmaceutical*”) AND (“*predicting*” OR “*prediction*”)”.

A execução da busca nas bases científicas considerou todos os artigos retornados nas seguintes bases, sendo: 2 artigos de IEEE Xplore, 22 de Springer Link e 47 de PubMed, totalizando 71 artigos. Os critérios de inclusão e exclusão estão detalhados na Tabela 1. A primeira avaliação considerou título, resumo e palavras-chave. Dois pesquisadores analisaram os artigos selecionados e elaboraram a lista final em consenso após discussões. Ao final os artigos relacionados a distribuição de fármacos foram identificados e discutidos com uma pesquisadora com experiência no domínio biomédico.

Tabela 1 – Critérios de inclusão e exclusão de artigos

Tipo	Sigla	Critério
Inclusão	I1	Pesquisas sobre ML, Farmacêutica e Predição.
	I2	Estudos que utilizam ML em distribuição de fármacos.
Exclusão	E1	Artigos escritos em idiomas diferentes do Inglês e do Português.
	E2	Artigos que não estejam relacionados com ML e distribuição de fármacos.
	E3	Artigos que não sejam da área de computação ou multidisciplinar com computação.
	E4	Textos que não sejam publicações científicas.
	E5	Resumos com menos de 4 páginas e que não tenham profundidade ou resultados relevantes.
	E6	Livros.

3. Soluções baseadas em ML para otimizar a Distribuição de Fármacos

Após a leitura completa, 6 artigos sobre soluções de ML para otimizar distribuição de fármacos foram selecionados e são apresentados nesta seção.

Deng et al. (2022) apresentam um modelo de predição para acelerar o desenvolvimento de produtos de microesferas para fármacos contendo moléculas pequenas por meio de técnicas de ML. Várias técnicas foram analisadas, tais como: *Deep Neural Network* (DNN), *Decision Tree* (DT) *K-Nearest Neighbors* (KNN), *Support Vector Machine* (SVM), *Partial Least Squares* (PLS), para citar algumas.

Lou e Hageman (2021) utilizam métodos de ML (*e.g.*, DT, ANN, KNN) para prever a biodisponibilidade (*i.e.*, velocidade e extensão da absorção do fármaco) após a administração subcutânea de anticorpos monoclonais, mesmo sem conhecer completamente o mecanismo e a causalidade entre entradas e saídas.

He et al. (2020) apresentam o uso de técnicas de ML (*e.g.*, KNN, PLS, DNN, SVM) para o desenvolvimento de formulações de nano cristais substituindo processos de tentativa e erro que consomem tempo e recursos e será de grande auxílio para os especialistas no processo.

Gao et al. (2021) propõem o uso de uma metodologia computacional integrada baseada em técnicas de ML (*e.g.*, RF, KNN, DT, SVM). O objetivo é diminuir os trabalhos tradicionais de design de formulações de medicamentos e trazer novas ideias para futuros projetos de formulações, modelagem molecular e abordagens experimentais para o design racional de formulações.

Noorain et al. (2023) abordam a combinação de nano-terapêuticos e a técnica de *Gaussian Process Models* (GPM) para simplificar os sistemas de desenvolvimento de medicamentos antivirais, automatizando a análise.

Por fim, Damiani & Damiani (2021) utilizam a combinação de tecnologias microfluídicas e de redes neurais artificiais, juntamente com o uso de biomateriais, para gerar micropartículas poliméricas carregadas com medicamentos.

4. Discussão e Conclusão

Em linha com o apresentado em He et al. (2020), a revisão aponta que o uso de ML na criação e seleção de formulações farmacêuticas, com o foco em distribuição de fármacos, tem potencial de apoiar a indústria farmacêutica. Tal constatação se deve à capacidade de acelerar os extensivos processos manuais de tentativa e erro. O atual cenário é bastante promissor para o uso dos algoritmos de ML, pois é crescente o volume de dados científicos [Kamerzell e Middaugh 2020] e isto possibilita a criação de modelos confiáveis. Entretanto, os estudos também apontam que os modelos preditivos devem ser usados de forma cuidadosa e complementar com a experimentação laboratorial, pois a validação experimental é essencial para confirmar as previsões dos modelos e garantir a qualidade das formulações.

Este artigo apresentou uma revisão de literatura sobre o uso de ML no contexto da criação e seleção de formulações farmacêuticas com foco em distribuição de fármacos. A abordagem empregada nesta revisão permitiu a verificação e análise de tendências, bem como abordagens tecnológicas adotadas ao longo dos últimos cinco anos. Em um universo de 71 artigos, 6 trabalhos foram criteriosamente selecionados, classificados e sintetizados de modo a representar o estado-da-arte das abordagens para avaliar e representar o uso de uma ferramenta computacional e sua aplicabilidade na área da saúde e no desenvolvimento de medicamentos com foco em distribuição de

fármacos. Foram apresentadas as técnicas e teorias utilizadas, os aspectos positivos e limitações dos estudos.

Pretende-se expandir esta revisão preliminar para outros aspectos ligados à formulação farmacêutica, bem como aprofundar nos aspectos científicos e tecnológicos de cada projeto avaliado.

Referências

- Bannigan P, Aldeghi M, Bao Z, Häse F, Aspuru-Guzik A, Allen C. Machine learning directed drug formulation development. *Adv Drug Deliv Rev.* 2021 Aug;175:113806. doi: 10.1016/j.addr.2021.05.016. Epub 2021 May 19. PMID: 34019959.
- Kamerzell TJ, Middaugh CR. Prediction Machines: Applied Machine Learning for Therapeutic Protein Design and Development. *J Pharm Sci.* 2021 Feb;110(2):665-681. doi: 10.1016/j.xphs.2020.11.034. Epub 2020 Dec 2. PMID: 33278409.
- Lou H, Hageman MJ. Machine Learning Attempts for Predicting Human Subcutaneous Bioavailability of Monoclonal Antibodies. *Pharm Res.* 2021 Mar;38(3):451-460. doi: 10.1007/s11095-021-03022-y. Epub 2021 Mar 12. PMID: 33710513.
- Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele University, 33(TR/SE-0401), 28. <https://doi.org/10.1.1.122.3308>
- Wang W, Ye Z, Gao H, Ouyang D. Computational pharmaceutics - A new paradigm of drug delivery. *J Control Release.* 2021 Oct 10;338:119-136. doi: 10.1016/j.jconrel.2021.08.030. Epub 2021 Aug 19. PMID: 34418520.
- Deng J, Ye Z, Zheng W, Chen J, Gao H, Wu Z, Chan G, Wang Y, Cao D, Wang Y, Lee SM, Ouyang D. Machine learning in accelerating microsphere formulation development. *Drug Deliv Transl Res.* 2023 Apr;13(4):966-982. doi: 10.1007/s13346-022-01253-z. Epub 2022 Dec 1. PMID: 36454434.
- He Y, Ye Z, Liu X, Wei Z, Qiu F, Li HF, Zheng Y, Ouyang D. Can machine learning predict drug nanocrystals? *J Control Release.* 2020 Jun 10;322:274-285. doi: 10.1016/j.jconrel.2020.03.043. Epub 2020 Mar 29. PMID: 32234511.
- Gao H, Jia H, Dong J, Yang X, Li H, Ouyang D. Integrated in silico formulation design of self-emulsifying drug delivery systems. *Acta Pharm Sin B.* 2021 Nov;11(11):3585-3594. doi: 10.1016/j.apsb.2021.04.017. Epub 2021 May 5. PMID: 34900538; PMCID: PMC8642610.
- Noorain L, Nguyen V, Kim HW, Nguyen LTB. A Machine Learning Approach for PLGA Nanoparticles in Antiviral Drug Delivery. *Pharmaceutics.* 2023 Feb 2;15(2):495. doi: 10.3390/pharmaceutics15020495. PMID: 36839817; PMCID: PMC9966002.
- Damiati SA, Damiati S. Microfluidic Synthesis of Indomethacin-Loaded PLGA Microparticles Optimized by Machine Learning. *Front Mol Biosci.* 2021 Sep 22;8:677547. doi: 10.3389/fmolb.2021.677547. PMID: 34631792; PMCID: PMC8493061.