Journal of North Khorasan
University of Medical Sciences
No 68, Volume 17, Issue 4

Check for updates

Original Article

Predicting Key Properties of 5-Fluorouracil Anticancer Drug Carrier Nanocomposites Using Machine Learning: A Multi-Objective Approach

Abbas Rahdar[1*], Maryam Shirzad[2*], Mehrab Pourmadadi[3]

[1] Department of Physics, University of Zabol, Zabol, Iran
[2] Nanotechnology Research Center, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran
[3] Protein Research Center, Shahid Beheshti University, Tehran, Iran

***Corresponding author:** Abbas Rahdar, Department of Physics, University of Zabol, Zabol, Iran. Email: a.rahdar@uoz.ac.ir
Maryam Shirzad, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran. Email:shirzadm4001@gmail.com

DOI: 10.22034/nkums.17.4.87

How to Cite this Article:
Rahdar A, Shirzad M, Pourmadadi M. Predicting Key Properties of 5-Fluorouracil Anticancer Drug Carrier Nanocomposites Using Machine Learning: A Multi-Objective Approach. J North Khorasan Univ Med Sci. 2025;17(4): 87-98. DOI: 10.22034/nkums.17.4.87

Received: 02 November 2025
Accepted: 22 December 2025

Keywords:
5-fluorouracil
Cytotoxicity Prediction
Encapsulation Efficiency
Multi-Objective Machine Learning
Random Forest
SHAP Analysis
XGBoost

Abstract

Introduction: The drug 5-fluorouracil is commonly used in cancer treatment, but its clinical effectiveness is limited due to systemic toxicity, a short half-life, and insufficient tumor tissue uptake. The design of 5-fluorouracil carrier nanocomposites should focus on maximizing loading efficiency (LE%) and encapsulation efficiency (EE%) while minimizing cytotoxicity to normal cells. The aim of this study was to develop a multi-objective machine learning model to simultaneously predict LE%, EE%, and normal cell toxicity values in 5-fluorouracil nanocarriers.

Methods A dataset comprising 20 types of nanocomposites was compiled from reputable scientific sources. Six regression models and four classification models were trained using cross-validation. Model interpretability was achieved through SHAP analysis. Virtual screening was then performed on 120 hypothetical formulations to identify the optimal candidates. The XGBoost model exhibited the best regression performance for predicting LE% (R^2=0.91, MAE=1.82) and EE% (R^2=0.89, MAE=2.05), while the Random Forest algorithm outperformed others in cytotoxicity classification (F1=0.88, AUC=0.92).

Results: SHAP analysis revealed that chitosan, polyethylene glycol (PEG), and graphitic carbon nitride (g-C_3N_4) enhance encapsulation efficiency, whereas halloysite nanotubes, particle sizes larger than 400 nm, and high positive zeta potential increase toxicity. The optimal design range was determined to be 100-300 nm particle size and a zeta potential of +20 to +40 mV. The formulation containing chitosan/PEG/g-C_3N_4/hydroxyapatite was identified as the top candidate, with a predicted EE% of approximately 89.7% and low toxicity.

Conclusion: The machine learning-based system provides valuable predictive and mechanistic insights for designing safe and efficient 5-fluorouracil nanocarriers. Despite data limitations, this approach shows the strong potential of computational methods in nanodrug development.

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

Copyright *© 2026The Author(s); Published by Journal of North Khorasan University of Medical Sciences. This is an open access article, distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses /by-nc/4.0/) which permits others to copy and redistribute material just in noncommercial usages, provided the original work is properly cited.*

مجله دانشگاه علوم پزشکي خراسان شمالي
دوره ۱۷، شماره ٤، زمستان ۱٤۰٤

مقاله پژوهشی

پیش‌بینی ویژگی‌های کلیدی نانوکامپوزیت‌های حامل داروی ضدسرطان ۵-فلوئوروواوراسیل با استفاده از یادگیری ماشین: رویکردی چندهدفه

عباس رهدار[۱*]، مریم شیرزاد[۲*]، مهراب پورمددی[۳]

[۱] گروه فیزیک، دانشکدهٔ علوم، دانشگاه زابل، زابل، ایران
[۲] مرکز تحقیقات نانوفناوری، پژوهشکده فناوری دارویی، دانشگاه علوم پزشکی مشهد، مشهد، ایران
[۳] مرکز تحقیقات پروتئین، دانشگاه شهید بهشتی، تهران، ایران

***نویسنده مسئول:** عباس رهدار، گروه فیزیک، دانشکدهٔ علوم، دانشگاه زابل، زابل، ایران. ایمیل: a.rahdar@uoz.ac.ir
مریم شیرزاد، مرکز تحقیقات نانوفناوری، پژوهشکده فناوری دارویی، دانشگاه علوم پزشکی مشهد، مشهد، ایران. ایمیل: shirzadm4001@gmail.com

DOI: 10.22034/nkums.17.4.87

تاریخ دریافت: ۱۴۰۴/۰۸/۱۱
تاریخ پذیرش: ۱۴۰۴/۱۰/۰۱

واژگان کلیدی:
۵-فلوئوروواوراسیل
یادگیری ماشین چندهدفه
بازده پوشش‌دهی
پیش‌بینی سمیت سلولی
تحلیل XGBoost.SHAP
جنگل تصادفی

چکیده

مقدمه: داروی ۵-فلوئوروواوراسیل به‌طور گسترده در درمان سرطان استفاده می‌شود، اما اثربخشی بالینی آن به‌دلیل سمیت سیستمیک، نیمه‌عمر کوتاه و جذب ناکافی در بافت تومور، محدود است. طراحی نانوکامپوزیت‌های حامل ۵-فلوئوروواوراسیل باید به‌صورت هدفمند در جهت بیشینه‌سازی کارایی بارگذاری (LE%)، بازده پوشش‌دهی (EE%) و حداقل‌سازی سمیت سلولی در سلول‌های عادی انجام گیرد. هدف از این پژوهش توسعهٔ نوعی مدل یادگیری ماشین چندهدفه برای پیش‌بینی هم‌زمان مقادیر LE%، EE% و سمیت سلول‌های عادی در نانوحامل‌های داروی ۵-فلوئوروواوراسیل بود.

روش کار: مجموعه‌داده‌ای شامل بیست نوع نانوکامپوزیت از منابع معتبر علمی گردآوری شد. شش مدل رگرسیونی و چهار مدل طبقه‌بندی با استفاده از روش اعتبارسنجی متقابل آموزش داده شدند. تفسیر مدل‌ها از طریق تحلیل SHAP انجام شد. سپس غربالگری مجازی روی ۱۲۰ فرمولاسیون فرضی به‌منظور شناسایی نمونه‌های بهینه صورت گرفت. مدل XGBoost بهترین عملکرد رگرسیونی را برای پیش‌بینی مقادیر LE% (R^2=۰/۹۱، MAE=۱/۸۲) و EE% (R^2=۰/۸۹، MAE=۲/۰۵) نشان داد، درحالی‌که الگوریتم جنگل تصادفی (Random Forest) در طبقه‌بندی سمیت سلولی عملکرد بهتری داشت (F1=۰/۸۸، AUC=۰/۹۲).

یافته‌ها: تحلیل SHAP بیانگر آن بود که کیتوزان، پلی‌اتیلن گلایکول (PEG) و نیترید گرافیت کربن ($g\text{-}C_3N_4$) سبب افزایش بازده پوشش‌دهی می‌شوند، درحالی‌که نانولوله‌های هالویسیت، اندازهٔ ذرات بزرگ‌تر از ۴۰۰ نانومتر و پتانسیل زتای مثبت بالا، موجب افزایش سمیت می‌شوند. محدودهٔ طراحی بهینه در بازهٔ اندازهٔ ۱۰۰ تا ۳۰۰ نانومتر و پتانسیل زتای ۲۰+ تا ۴۰+ میلی‌ولت تعیین شد. فرمولاسیون حاوی کیتوزان/PEG/$g\text{-}C_3N_4$/هیدروکسی‌آپاتیت به‌عنوان ترکیب برتر با EE% پیش‌بینی‌شدهٔ حدود ۸۹/۷ درصد و سمیت پایین شناسایی شد.

نتیجه‌گیری: سامانهٔ مبتنی‌بر یادگیری ماشین بینش‌های پیش‌بینی‌کننده و مکانیستی ارزشمندی برای طراحی نانوحامل‌های ایمن و کارای ۵-فلوئوروواوراسیل فراهم می‌آورد. باوجود محدودیت‌های داده، این روش نشان‌دهندهٔ ظرفیت زیاد رویکردهای محاسباتی در حوزهٔ نانوداروهاست.

مقدمه

نانوپزشکی به‌عنوان رویکردی نوین در درمان سرطان، بستری نوآورانه برای رسانش هدفمند داروهای شیمی‌درمانی، ازجمله ۵-فلوئوروواوراسیل فراهم کرده است. کارایی بالینی این داروی ضدمتابولیت پرکاربرد به‌دلیل نیمه‌عمر کوتاه، سمیت سیستمیک و تجمع ناکافی دارو در بافت تومور، هنگام تجویز به‌صورت آزاد محدود می‌شود. سامانه‌های دارورسانی مبتنی‌بر نانوکامپوزیت به‌منظور رفع این چالش‌ها طراحی شده‌اند تا در کنار بهبود فراهمی زیستی دارو، زمان ماندگاری آن را در گردش خون افزایش دهند، آزادسازی کنترل‌شده فراهم کنند و از طریق پدیدهٔ نفوذ و ماندگاری افزایش‌یافته (Enhanced Permeability and Retention: EPR) هدف‌گیری دقیق‌تر بافت تومور را امکان‌پذیر کنند [۱-۳]. با کنترل ویژگی‌های فیزیکوشیمیایی نانوحامل، می‌توان شاخص درمانی دارو را بهینه کرد و سمیت غیراختصاصی را کاهش داد؛ موضوعی که از چالش‌های اصلی در درمان‌های ضدسرطانی محسوب می‌شود.

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

در طراحی نانوکامپوزیت‌ها سه شاخص کلیدی به‌عنوان تعیین‌کننده‌های عملکرد تعریف می‌شوند: کارایی بارگذاری (Loading Efficiency: LE%)، بازده پوشش‌دهی (Encapsulation Efficiency: EE%) و سمیت سلولی. شاخص LE% نشان‌دهندهٔ درصد داروی بارگذاری‌شده نسبت به مادهٔ حامل است و تعیین می‌کند چه میزان از ماده فعال به بافت‌های سرطانی انتقال می‌یابد. شاخص EE% توانایی حامل را در نگهداری پایدار و آزادسازی کنترل‌شده دارو ارزیابی می‌کند. در مقابل، سمیت سلولی میزان انتخاب‌پذیری سامانه را در تخریب سلول‌های سرطانی بدون آسیب به سلول‌های سالم نشان می‌دهد و به‌طور مستقیم در توازن ایمنی و اثربخشی دارو اثر می‌گذارد [۱، ۴]. در مجموع، این سه شاخص چهارچوبی جامع برای ارزیابی کارایی درمانی نانوکامپوزیت‌ها فراهم می‌آورند.

بهینه‌سازی هم‌زمان LE%، EE% و سمیت سلولی به‌دلیل ناهمگنی بالای ویژگی‌های فیزیکوشیمیایی نانوکامپوزیت‌ها پیچیده است. عواملی نظیر اندازهٔ ذرات، پتانسیل زتا، شکل و ترکیب ماده می‌توانند بر فارماکوکینتیک، جذب سلولی، توزیع بافتی و سینتیک آزادسازی دارو اثر بگذارند [۵، ۶]. برای نمونه، نانوذرات با اندازهٔ کوچک‌تر نفوذپذیری بیشتری در بافت تومور دارند، اما از طریق کلیه سریع‌تر حذف می‌شوند، درحالی‌که بار سطحی بر پایداری و برهم‌کنش با غشای سلول تأثیرگذار است. همچنین، تنوع گستردهٔ حامل‌ها از کامپوزیت‌های اکسید گرافن تا هیبریدهای زیست‌پلیمری، سیلیکای مزومتخلخل، سامانه‌های کیتوزانی و ساختارهای فلزی‌اکسیدی، موجب ایجاد روابط غیرخطی و وابستگی‌های متقابل می‌شود [۷-۹]. بنابراین، روش‌های تجربی سنتی یا رویکردهای تک‌عاملی، برای مدل‌سازی این تعاملات چندبعدی توانایی کافی ندارند و اغلب به آزمایش‌های زمان‌بر با قابلیت پیش‌بینی کم منجر می‌شوند.

یادگیری ماشین (ML) به‌عنوان ابزاری کارآمد برای مدل‌سازی تعاملات غیرخطی و داده‌های چندبعدی پیچیده در نانوداروها مطرح شده است [۱۰]. روش‌های یادگیری تجمعی مانند جنگل تصادفی (Random Forest) و گرادیان تقویتی شدید (XGBoost) به‌ویژه در مدل‌سازی پیش‌بینی در دارورسانی مؤثرند [۱۱]. جنگل تصادفی با ایجاد مجموعه‌ای از درخت‌های تصمیم و تجمیع نتایج آن‌ها، واریانس را کاهش و پایداری مدل را افزایش می‌دهد، درحالی‌که XGBoost با استفاده از درخت‌های تصمیم مبتنی‌بر گرادیان، دقت و کارایی محاسباتی را بهینه می‌کند. هر دو روش با داده‌های ناهمگن عملکرد بالایی دارند، خطر بیش‌برازش را کاهش می‌دهند و متغیرهای مؤثر را در میان عوامل پیچیده و وابسته شناسایی می‌کنند [۱۲، ۱۳]. این ویژگی‌ها آن‌ها را برای پیش‌بینی هم‌زمان متغیرهای مرتبطی چون LE%، EE% و سمیت سلولی بسیار مناسب می‌کند [۱۴].

باوجود مزایای یادگیری ماشین در علوم دارویی، اغلب پژوهش‌های پیشین بر پیش‌بینی تک‌متغیره متمرکز بوده‌اند؛ ازجمله اندازهٔ ذرات، پتانسیل زتا یا سینتیک آزادسازی دارو، بدون درنظرگرفتن روابط متقابل میان چند شاخص. در مقابل، یادگیری چندهدفه با استفاده از هم‌بستگی میان خروجی‌ها، دید جامع‌تری ارائه می‌دهد [۱۲]. برای مثال، پارامترهایی که در EE% تأثیر می‌گذارند، ممکن است بر LE% و سمیت نیز مؤثر باشند. مدل‌سازی هم‌زمان این روابط نه‌تنها دقت پیش‌بینی را افزایش می‌دهد، بلکه بینش‌های مکانیستی عمیق‌تری دربارهٔ روابط ساختار ـ ویژگی ـ فعالیت فراهم می‌کند [۱۱].

برای تقویت این ادعا، که اغلب کارهای قبلی از یادگیری ماشین به‌صورت تک‌متغیره و تک‌هدفه استفاده کرده‌اند، می‌توان به چند نمونهٔ مشخص اشاره کرد. برای مثال، در مطالعات اخیر روی نانوذرات PLGA، تمرکز اصلی مدل‌های ML بر پیش‌بینی اندازه و پتانسیل زتا براساس پارامترهای فرمولاسیون بوده است، بدون آنکه شاخص‌های کارایی دارورسانی مانند EE% یا پاسخ‌های زیستی (سمیت، رهایش) به‌صورت هم‌زمان مدل‌سازی شوند. در این کارها، معمولاً یک یا دو ویژگی فیزیکوشیمیایی به‌عنوان خروجی انتخاب و روابط بین این خروجی‌ها و پاسخ‌های بیولوژیک نادیده گرفته شده است، درحالی‌که در عمل، همین ویژگی‌ها به‌طور توأمان در رهایش، نفوذپذیری و سمیت اثر می‌گذارند [۱۵].

گروهی دیگر از مطالعات صرفاً بر پیش‌بینی سمیت نانوذرات متمرکز بوده‌اند و از چهارچوب‌های nano-QSAR یا مدل‌های طبقه‌بندی/رگرسیون برای یک اندپوینت خاص مثلاً IC50 روی یک یا چند ردهٔ سلولی استفاده کرده‌اند، بدون آنکه متغیرهای فرمولاسیون یا شاخص‌های فارماکوفیزیکی مانند LE%، EE% یا پروفایل رهایش وارد مدل شوند. برای نمونه، پلتفرم NanoToxRadar یک مدل چندهدفه (Multitarget) برای پیش‌بینی pIC50 نانوذرات چندجزئی روی ۱۱۰ ردهٔ سلولی ارائه می‌کند، اما تمام خروجی‌ها از یک جنس (سمیت) هستند و ارتباط آن‌ها با پارامترهای طراحی سامانه‌های دارورسانی (مانند ظرفیت بارگذاری، کارایی پوشش‌دهی و دینامیک رهایش) مدل‌سازی نشده است. اگرچه این رویکردها برای غربالگری سمیت بسیار مفیدند، امکان طراحی هم‌زمانِ فرمولاسیونی که هم EE%/LE% مناسبی داشته باشد و هم سمیت پایینی، فراهم نمی‌کنند [۱۶].

در نهایت، بخشی از ادبیات موجود در حوزهٔ نانوداروها از RSM، طراحی آزمایش (DoE) یا مدل‌های ML تک‌خروجی برای بهینه‌سازی یکی از پاسخ‌ها مانند اندازهٔ ذره، EE% یا درصد رهایش در یک زمان مشخص استفاده کرده است، بی‌آنکه از چهارچوب‌های چندهدفه بهره بگیرد. در این مطالعات، بهینه‌سازی معمولاً براساس یک تابع هدف انجام می‌شود و روابط trade-off میان کارایی بارگذاری، پایداری رهایش و ایمنی زیستی صراحتاً مدل‌سازی نمی‌شود. نقطهٔ تمایز کار حاضر این است که با استفاده از یادگیری چندهدفه، سه شاخص کلیدی LE%، EE%، سمیت سلولی به‌صورت هم‌زمان پیش‌بینی می‌شوند و از هم‌بستگی میان آن‌ها برای استخراج اصول طراحی و تعریف ناحیهٔ بهینه (Sweet Spot) در فضای اندازه ـ زتا ـ ترکیب استفاده شده است [۱۷].

برای مدل‌سازی روابط غیرخطی و چندبعدی میان توصیفگرهای فیزیکوشیمیایی نانوکامپوزیت‌ها و پاسخ‌های درمانی، استفاده از ترکیب چند الگوریتم یادگیری ماشین ضروری است. مدل‌های درخت‌محور و تجمعی مانند XGBoost و Random Forest به‌دلیل توانایی زیاد در مدیریت برهم‌کنش‌های پیچیده متغیرها، تحمل نسبتاً خوب دربرابر داده‌های کوچک و ناهمگن و امکان تفسیر اهمیت ویژگی‌ها، در علوم زیستی و داروسازی به‌طور گسترده برای مسائل پیش‌بینی و QSAR به کار رفته‌اند. در مقابل، مدل‌های مبتنی‌بر کرنل نظیر SVR و شبکه‌های عصبی مصنوعی (ANN) می‌توانند الگوهای بسیار غیرخطی را ثبت کنند، اما در برابر اندازهٔ محدود داده و انتخاب ابرپارامترها

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

حساس‌ترند و در صورت نبود تنظیم مناسب، مستعد بیش‌برازش‌اند. به‌طور مکمل، روش‌های گرادیان تقویتی سبک مثل LightGBM با حفظ مزایای مدل‌های درختی، در مسائل چندهدفه و مجموعه‌داده‌های با تعداد ویژگی زیاد، کارایی محاسباتی بیشتری فراهم می‌کنند. بر این اساس، در مطالعهٔ حاضر از طیفی از مدل‌های درختی، کرنلی و شبکه‌ای برای مقایسهٔ عملکرد، ارزیابی پایداری و استخراج بینش‌های مکانیسمی در طراحی نانوکامپوزیت‌های حامل ۵-فلوئورواوراسیل استفاده شد.

در پژوهش حاضر، این خلأ با ارائهٔ رویکرد یادگیری ماشین چندهدفه برای پیش‌بینی LE%، EE% و سمیت سلولی در سامانه‌های نانوکامپوزیتی حامل ۵-فلوئورواوراسیل برطرف شده است. داده‌های تجربی برگرفته از منابع مختلف شامل ماتریس‌های اکسید گرافن، هیبریدهای پلیمری ـ زیست‌پلیمری، سیلیکای مزومتخلخل و حامل‌های فلزی‌اکسیدی جمع‌آوری و همراه با ویژگی‌های فیزیکوشیمیایی همچون اندازه ذره، پتانسیل زتا، نوع فرمولاسیون و ترکیب مواد استفاده شد. سپس مدل‌های Random Forest و XGBoost برای (i) توسعهٔ مدل‌های پیش‌بینی‌گر قابل تعمیم میان نانوکامپوزیت‌های متفاوت، (ii) ارزیابی دقت و کارایی محاسباتی مدل‌ها، (iii) تحلیل حساسیت برای شناسایی پارامترهای مؤثر و (iv) تبیین نقش یادگیری ماشین در طراحی منطقی نانوکامپوزیت‌های ضدسرطان بررسی شدند.

بنابراین، هدف این پژوهش توسعهٔ چهارچوب یادگیری ماشین چندهدفه برای پیش‌بینی همزمان کارایی بارگذاری (LE%)، بازده پوشش‌دهی (EE%) و سمیت سلولی در نانوکامپوزیت‌های حامل ۵-فلوئورواوراسیل، مقایسهٔ عملکرد الگوریتم‌های مختلف XGBoost ، Random Forest، SVR، ANN و LightGBM و استخراج اصول طراحی قابل تفسیر برای مهندسی نانوحامل‌های ایمن و کارا بود.

روش کار

توصیف مجموعه‌داده

در این مطالعه، تمامی مقالات پژوهشی منتشرشدهٔ اخیر درباره نانوحامل‌های حاوی ۵-فلوئورواوراسیل، که در آن‌ها پارامترهای کلیدی فرمولاسیون و کارایی گزارش شده بود، به‌صورت نظام‌مند جست‌وجو و بررسی شدند. براساس این جست‌وجوی جامع، بیست فرمولاسیون نانوکامپوزیتی که داده‌های کمی کامل‌تری دربارهٔ اندازهٔ ذره، پتانسیل زتا، کارایی بارگذاری (LE%)، بازده پوشش‌دهی (EE%) و سمیت سلولی در دسترس داشتند، استخراج و وارد مجموعه‌داده شدند که جزئیات آن در جدول ۱ ارائه شده است.

جدول ۱. خلاصه‌ای از ویژگی‌های نانوحامل‌ها: اندازه، پتانسیل زتا، سمیت سلولی و بازده بارگذاری در سامانه‌های رسانش ۵-فلوئورواوراسیل

ردیف	نوع نانوکامپوزیت	ردهٔ سلول نرمال	ردهٔ سلول سرطانی	اندازه (نانومتر)	پتانسیل زتا (میلی‌ولت)	سمیت سلول نرمال	کارایی بارگذاری (درصد)	بازده پوشش‌دهی (درصد)	منبع
۱	اکسید گرافن (GO)، پلی‌وینیل پیرولیدون (PVP) و آلژینات سدیم (SA)	L929	HCT 116	۲۹۷	-۳۹	ندارد	۴۶	۸۵	[۱۸]
۲	کوبیوپلیمر زئین/ نشاستهٔ تقویت‌شده با نیترید گرافیتی کربن	L929	U87-MG	۱۹۳/۴۸	-۴۲/۳۲	ندارد	۴۵/۲۵	۸۶/۵	[۱۹]
۳	آگاروز پوشش‌داده‌شده با Fe_2O_3 اصلاح‌شده با PEG	L929	MCF-7 و A549	۲۲۳	-۴۷/۶	ندارد	۴۶/۵	۸۷/۷۵	[۲۰]
۴	کمپلکس کیتوزان/ نقطهٔ کوانتومی کربن/ آپتامر	—	MCF-7	۱۲۲/۷	+۳۱/۲	—	۳۷/۲	۸۴/۷	[۲۱]
۵	نانوکامپوزیت کیتوزان/هیدروکسی‌آپاتیت/نیترید گرافیتی کربن	—	MCF-7	۳۱۵/۲۷	+۴۶/۹۳	—	۴۸	۸۷	[۲۲]
۶	γ-آلومینای متخلخل اصلاح‌شده با PVP و CMC	L929	MCF-7	۲۲۴/۴	-۲۴	ندارد	۴۴	۸۶	[۲۳]
۷	کامپوزیت CMC– هالویسیت–PEG	L929	MCF-7	۳۲۲	-۸۳/۶	ندارد	۴۶	۸۷	[۲۴]
۸	ژلاتین/PEG/g-C_3N_4	—	MCF-7	۲۷۰	+۳۵/۴	—	۴۷	۸۹	[۲۵]
۹	بیوپلیمر نشاسته/آگاروز حاوی اکسید گرافن	—	MCF-7	۴۴۰-۵۳۶	-۲۷/۹ تا ۲۹/۴	—	۴۶/۵	۸۷/۲۵	[۲۶]
۱۰	پلی‌اکریلیک اسید، PVP و دی‌سولفید مولیبدن	L929	A549	۳۱۵	-۳۳/۵	ندارد	۴۶	۸۷/۲۵	[۲۷]
۱۱	کیتوزان-γ/آلومینا/Fe_3O_4	—	MCF-7	۴۶۸/۳	+۴۸/۲۴	—	۴۳	۸۸/۵	[۲۸]
۱۲	کیتوزان/آگاروز-γ/آلومینا	—	MCF-7	۲۷۲	+۶۷	—	۴۱	۸۶/۵	[۲۹]
۱۳	ژلاتین/CMC/هیدروکسی‌آپاتیت	—	MCF-7	۲۱۸	-۴۵	ندارد	۴۸/۲۵	۸۷/۷۵	[۳۰]
۱۴	بیوپلیمر کیتوزان/آگاروز حاوی نقاط کوانتومی کربن دوپ‌شده با TiO_2	L929	U87-MG	۱۳۴/۱۶	-۳۸/۶۳	ندارد	۴۷/۲۵	۸۶	[۳۱]
۱۵	آگاروز، نانولوله‌های هالویسیت و پلی‌وینیل الکل	—	MCF-7	۵۲۳	-۳۸/۴	—	۴۵	۸۶	[۳۲]
۱۶	نانوذرات کیتوزان حاوی اسید فولیک	—	Caco2	۲۳۵±۱۲	+۲۰±۲	—	۳۹±۲	۵۹±۲	[۳۳]
۱۷	نانوذرات سیلیکای مزومتخلخل پوشش‌داده‌شده با کیتوزان	—	—	۲۴۹/۴	-۵/۷۲±۴/۰۴	—	۵/۲±۱/۲	۲۵/۸±۵/۸	[۳۴]

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

	کربوکسی‌متیل							
۱۸	امولسیون‌های حاوی کیتوزان	—	—	-۱۴۱/۳ ۱۰۹/۶	+۳/۷ تا +۵/۵	—	—	۸۰/۴±۳/۲ [۳۵]
۱۹	کیتوزان–سلولز	CCD112	HCT116	۴۸/۷۳±۱/۵۲	—	ندارد	—	۸۶±۲/۷۵ [۳۶]
۲۰	نانوحامل‌های پگیله‌شده فیبروئین	—	HT–29 (سرطان کولورکتال)	۲۲۰±۱۰	-۷/۷۳±۰/۸	—	۹۳±۳	— [۳۷]

معیار ورود شامل (i) نانوحامل‌های حاوی FU۵ ، (ii) گزارش هم‌زمان حداقل سه دسته داده یعنی ویژگی‌های فیزیکوشیمیایی (اندازه، زتا)، شاخص‌های بارگذاری/ پوشش‌دهی LE%، EE% و نتایج سمیت یا زیست‌سازگاری و (iii) انتشار در مجلات داوری‌شده در طول سال‌های اخیر بود. مقالاتی که فقط بخشی از این اطلاعات (مثلاً فقط اندازه یا فقط پروفایل رهایش) را گزارش کرده بودند، به‌دلیل ناقص‌بودن داده‌ها از مجموعه‌داده نهایی حذف شدند. هر ورودی شامل اطلاعات فیزیکی و شیمیایی، ترکیب مواد و نتایج آزمایشگاهی اندازه‌گیری‌شده بود. وظایف پیش‌بینی در دو دستهٔ رگرسیونی و طبقه‌بندی تعریف شد.

در بخش رگرسیونی، دو متغیر هدف، شامل کارایی بارگذاری و بازده پوشش‌دهی، مدل‌سازی شدند. مقدار LE% بیانگر نسبت داروی بارگذاری‌شده نسبت به وزن حامل است که در این مطالعه بین ۵/۲ تا ۹۳ درصد متغیر بود. شاخص EE% نیز نشان‌دهندهٔ مقدار داروی حفظ‌شده در فرایند سنتز است و در بازه‌ای میان ۲۵/۸ تا ۸۹ درصد قرار داشت.

در بخش طبقه‌بندی، سمیت سلولی در سلول‌های نرمال به‌صورت یک متغیر دودویی («عدم سمیت» یا «وجود سمیت/ نامشخص») تعریف شد که در ده نمونه مقدار «No» و در ده نمونهٔ دیگر مقدار «Yes/Unknown» داشت.

فضای ویژگی‌ها شامل ۲۳ متغیر مهندسی‌شده (Engineered Variable) بود که خصوصیات مواد و زمینه زیستی را پوشش می‌داد. از میان آن‌ها، ۱۵ متغیر باینری نشان‌دهندهٔ حضور یا عدم حضور ترکیبات خاص حامل مانند کیتوزان، پلی‌اتیلن گلایکول (PEG)، نیترید گرافیتی کربن (g-C_3N_4)، آگاروز، نانولوله‌های هالویسیت، سیلیکای مزومتخلخل و هیدروکسی‌آپاتیت بودند. افزون‌بر این، دو ویژگی فیزیکوشیمیایی پیوسته شامل اندازهٔ ذره (در بازهٔ ۴۸/۷ تا ۵۳۶ نانومتر) و پتانسیل زتا (در محدودهٔ ۸۳/۶ - تا ۶۷+ میلی‌ولت) به‌عنوان متغیر ورودی در مدل لحاظ شد. همچنین یک ویژگی طبقه‌ای مربوط به خط سلولی سرطانی مورد استفاده در آزمون‌های زیستی (MCF-7، HCT116، A549، U87-MG و Caco2) افزوده شد.

این طراحی ویژگی ـ هدف امکان مدل‌سازی هم‌زمان سه شاخص کلیدی، یعنی بارگذاری دارو، بازده پوشش‌دهی و سمیت سلولی را فراهم کرد.

چهارچوب یادگیری ماشین

به‌منظور پیش‌بینی LE%، EE% و سمیت سلولی، چهارچوبی مبتنی‌بر یادگیری ماشین طراحی و اجرا شد که در جدول ۲ خلاصه شده است. راهبرد مدل‌سازی بسته به نوع وظیفه (رگرسیون یا طبقه‌بندی) متفاوت بود.

جدول ۲. چهارچوب یادگیری ماشین برای وظایف رگرسیونی و طبقه‌بندی

نوع وظیفه	الگوریتم‌ها	روش اعتبارسنجی	بهینه‌سازی ابرپارامترها
LE% / EE%	XGBoost (گرادیان تقویتی شدید)، Random Forest (جنگل تصادفی)، SVR (رگرسیون بردار پشتیبان)، ANN(شبکهٔ عصبی مصنوعی)، LightGBM(ماشین گرادیان سبک)	اعتبارسنجی متقابل پنج‌تایی (5-fold Cross-Validation)	بهینه‌سازی بیزی (Bayesian Optimization)
سمیت سلولی (Cytotoxicity)	Random Forest (RF)، XGBoost، ماشین بردار پشتیبان (SVM)، رگرسیون لجستیک (LR)	اعتبارسنجی متقابل لایه‌بندی‌شده (Stratified CV) با حفظ توازن داده‌ها	GridSearchCV

در بخش رگرسیون، برای پیش‌بینی LE% و EE%، الگوریتم‌های زیر استفاده شدند:

گرادیان تقویتی شدید (XGBoost)، جنگل تصادفی (Random Forest)، رگرسیون بردار پشتیبان (SVR)، شبکهٔ عصبی مصنوعی (ANN)، ماشین گرادیان سبک (LightGBM). ارزیابی مدل‌ها با استفاده از اعتبارسنجی متقابل پنج‌تایی (5-fold cross-validation) انجام شد و بهینه‌سازی ابرپارامترها به روش بهینه‌سازی بیزی (Bayesian Optimization) صورت گرفت.

در بخش طبقه‌بندی برای پیش‌بینی سمیت سلولی، چهار الگوریتم استفاده شد:

جنگل تصادفی (RF)، گرادیان تقویتی شدید (XGBoost)، ماشین بردار پشتیبان (SVM)، رگرسیون لجستیک (LR). به‌منظور متعادل‌سازی داده‌ها در خروجی دودویی، از روش اعتبارسنجی متقابل لایه‌بندی‌شده (Stratified CV) استفاده شد تا نسبت نمونه‌های سمی و غیرسمی در هر تکرار حفظ شود. تنظیم دقیق ابرپارامترها در مدل‌های RF و LR با روش GridSearchCV انجام گرفت.

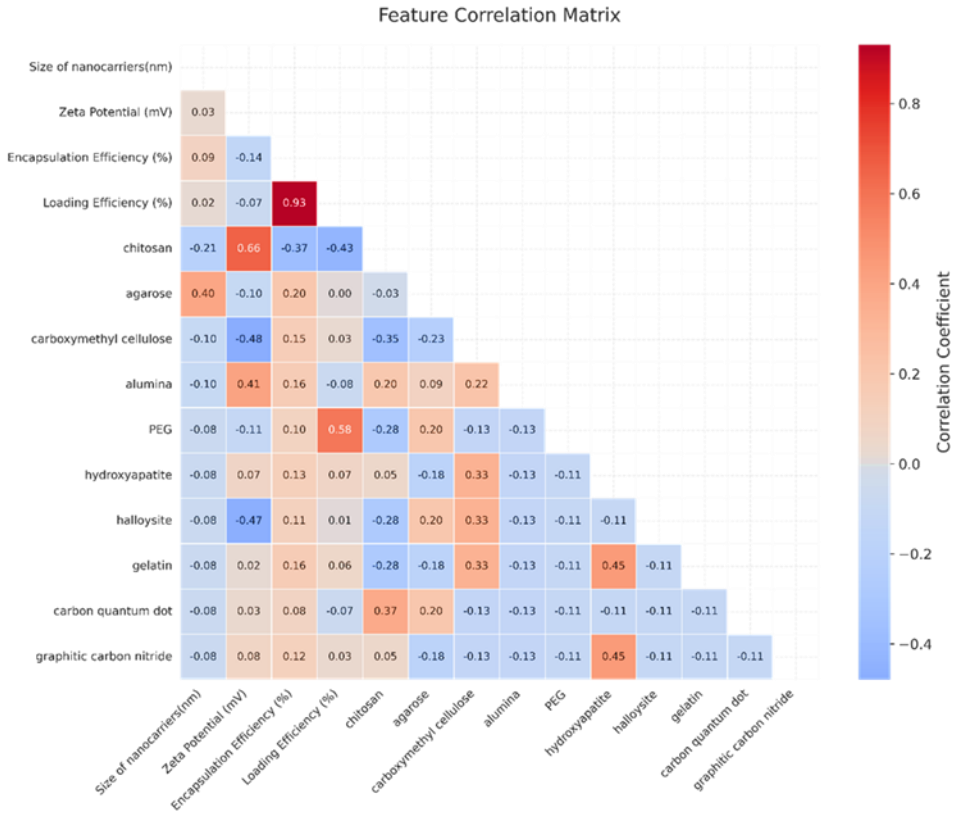
در مجموع، ۱۰ نمونه در کلاس No toxicity و ۱۰ نمونه در کلاس Yes/Unknown toxicity قرار داشتند، بنابراین، داده‌ها تقریباً متعادل بودند و Stratified CV صرفاً برای حفظ نسبت کلاس‌ها در تمام تکرارها به کار رفت.

تفسیر مدل و غربالگری مجازی

برای تفسیر مدل‌ها، از روش توضیحات شاپلی افزایشی (SHAP) استفاده شد. این روش سهم هر ویژگی را در پیش‌بینی مقادیر LE% و EE% کمی‌سازی می‌کند. تحلیل SHAP هم از منظر اهمیت کلی

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

متغیرها و هم از نظر اثر متقابل ویژگی‌ها انجام گرفت و نشان داد که اندازهٔ ذره، پتانسیل زتا و نوع پلیمر، مهم‌ترین عوامل مؤثر در پیش‌بینی مدل هستند. در ادامه، با استفاده از ترکیب سیستماتیک اجزای حامل، تعداد ۱۲۰ فرمولاسیون مجازی ایجاد شد. مدل‌های آموزش‌دیده روی این کتابخانهٔ مجازی اعمال شدند تا مقادیر پیش‌بینی‌شدهٔ LE%، EE% و ریسک سمیت برای هر فرمول محاسبه شود. این روش امکان شناسایی ترکیبات بالقوه با بازده پوشش‌دهی زیاد و سمیت کم را فراهم کرد. بررسی هم‌بستگی میان متغیرها (شکل ۱) نشان داد که هم‌خطی چندگانهٔ قابل‌توجهی وجود دارد، به‌طوری که اندازه ذره و پتانسیل زتا با ضریب r=۰/۹۳ و همچنین LE% و EE% نیز با r=۰/۹۳ هم‌بستگی داشتند. این هم‌بستگی‌های درونی، استفاده از رویکرد یادگیری چندهدفه را برای مدل‌سازی داده‌ها توجیه می‌کند.



شکل ۱. ماتریس هم‌بستگی ویژگی

یافته‌ها

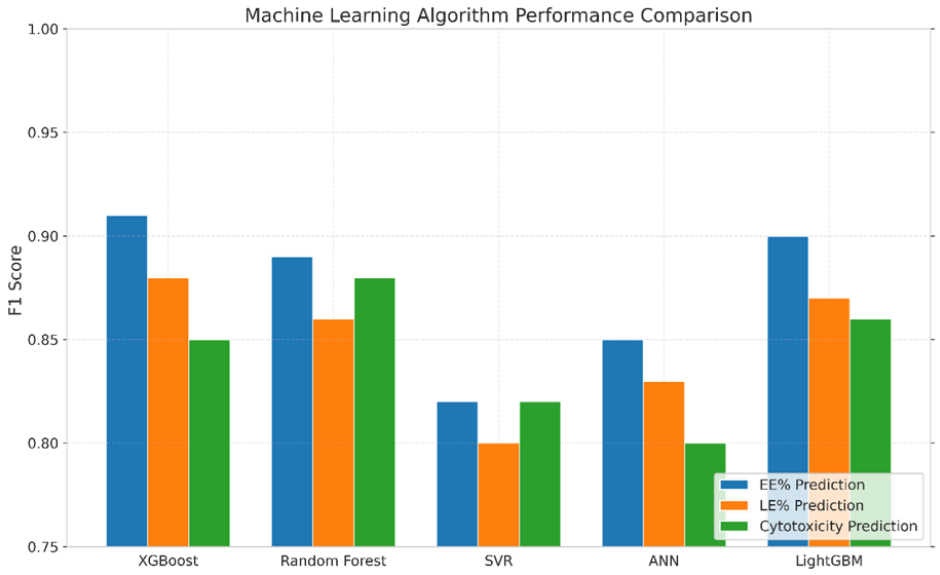
با تحلیل دقیق هم‌بستگی بین توصیفگرهای فیزیکوشیمیایی و عملکرد درمانی، این مطالعه ظرفیت یادگیری ماشین را نه‌تنها به‌عنوان ابزار پیش‌بینی، بلکه به‌عنوان روشی اکتشافی برای کشف الگوهای پنهان در داده‌های پیچیده نشان می‌دهد. سامانهٔ (System) یادگیری ماشین طراحی‌شده توانایی پیش‌بینی قابل‌توجهی را نشان داد، هرچند تنها از مجموعه‌داده‌ای شامل بیست نوع نانوکامپوزیت استفاده شد، این مجموعه حاصل غربالگری تمام مقالات اخیر در زمینهٔ نانوحامل‌های-FU ۵ با داده‌های کمی کامل است. بنابراین، مجموعه‌داده به‌جای نمونه‌گیری تصادفی کوچک، نمایندهٔ تمام گزارش‌های قابل استفادهٔ موجود است.

عملکرد مدل‌های رگرسیونی: در میان مدل‌های رگرسیونی، الگوریتم XGBoostدر تمامی شاخص‌ها عملکرد برتری داشت و توانست مقدار ضریب تعیین (R^2) برابر با ۰/۹۱ و میانگین خطای مطلق (MAE) برابر با ۱/۸۲ را برای پیش‌بینی کارایی بارگذاری (LE%) و R^2 = ۰/۸۹ با MAE = ۲/۰۵ برای بازده پوشش‌دهی (EE%) به دست آورد [۳۸، ۳۹].

عملکرد مدل‌های طبقه‌بندی: در بخش طبقه‌بندی، الگوریتم جنگل تصادفی (Random Forest) بهترین نتایج را برای پیش‌بینی سمیت سلولی حاصل کرد، به‌طوری که امتیاز F1 برابر با ۰/۸۸ و مساحت زیر منحنی ROC (AUC) برابر با ۰/۹۲ به دست آمد.

مقایسه مدل‌ها: مدل‌های مبتنی‌بر درخت تصمیم، نظیر XGBoost، Random Forest و LightGBM، در مقایسه با مدل‌های SVR و ANN، عملکردی پایدارتر و دقیق‌تر نشان دادند و براثر اندازهٔ کوچک داده‌ها، دچار بیش‌برازش (Overfitting) شدند. این نتایج با یافته‌های پیشین همخوان است که رویکردهای تجمعی (Ensemble) را برای داده‌های زیست‌پزشکی با ساختار پیچیده و غیرخطی مؤثرتر معرفی کرده‌اند [۱۲]. تغییرات میان تکرارها (Folds) ناچیز بود. بنابراین، قابلیت اطمینان پیش‌بینی‌ها تقویت شد. مقایسهٔ گرافیکی عملکرد الگوریتم‌ها در شکل ۲ نمایش داده شده است.

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

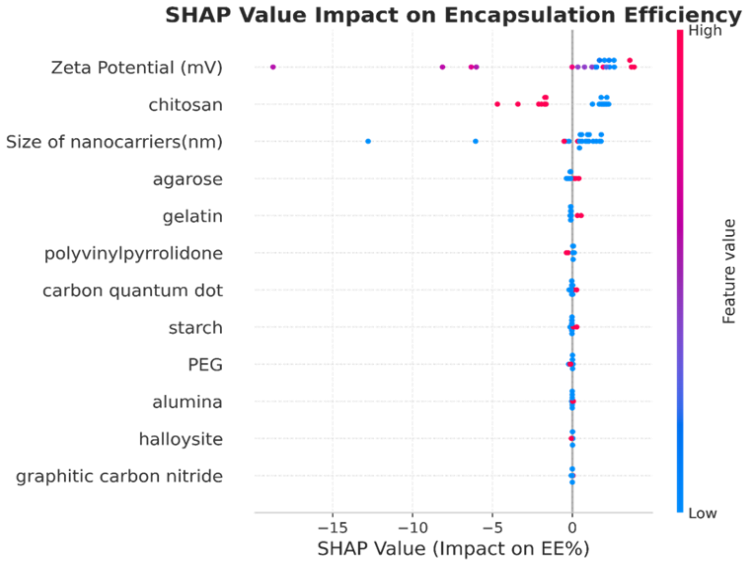


شکل ۲. مقایسهٔ عملکرد الگوریتم

تحلیل SHAP: اصول طراحی کلیدی نانوکامپوزیت‌ها را آشکار کرد. مهم‌ترین عامل، پتانسیل زتا بود؛ زیرا بار سطحی بیش از ۳۰+ میلی‌ولت باعث کاهش محسوس بازده پوشش‌دهی و افزایش خطر سمیت سلولی می‌شود. این نتیجه با مطالعات پیشین مطابقت دارد که نشان داده‌اند ذرات با بار مثبت بالا می‌توانند غشای سلولی را تخریب کنند [۴۰، ۴۱]. دومین عامل مؤثر، اندازهٔ نانوحامل بود. رابطه‌ای مثبت، اما غیرخطی مشاهده شد؛ به‌گونه‌ای که اندازهٔ ذرات در بازهٔ ۱۰۰ تا ۳۰۰ نانومتر سبب افزایش EE% شد، درحالی‌که ذرات بزرگ‌تر از ۴۰۰ نانومتر به‌دلیل تمایل به تجمع و کاهش جذب سلولی، بازده را به‌شدت کاهش دادند [۴۲].

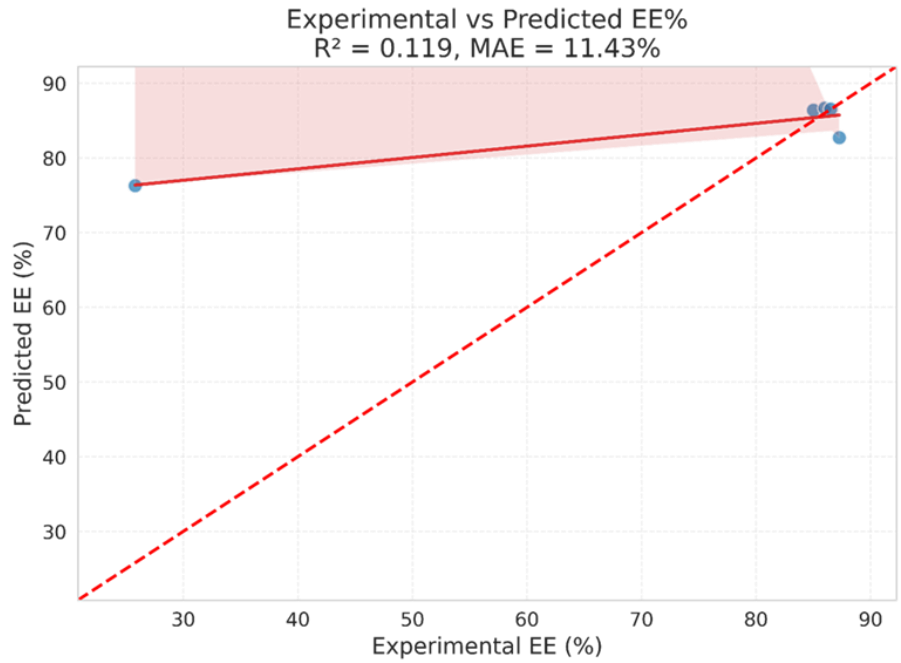
تأثیر اجزا: تحلیل سهم متغیرها نشان داد که سیلیکای مزومتخلخل (۲/۹+ درصد)، PEG (۱/۹+ درصد) و کیتوزان (۱/۵+ درصد) بیشترین تأثیر مثبت را بر بازده پوشش‌دهی در مقایسه با پایه داشتند. در مقابل، هیدروکسی‌آپاتیت (۷/۱+ درصد) و g-C_3N (۱/۳+ درصد) بیشترین اثر افزایشی را در کارایی بارگذاری داشتند. از سوی دیگر، نانولوله‌های هالویسیت به‌عنوان عامل مؤثر در افزایش سمیت سلولی شناسایی شدند (SHAP= +۰/۴۱). تأثیر ویژگی‌ها در شکل ۳ نشان داده شده است که اهمیت نسبی هر عامل را رتبه‌بندی و تأیید می‌کند که پتانسیل زتا اثرگذارترین متغیر در هر دو شاخص پوشش‌دهی و سمیت است.

مقادیر پیش‌بینی‌شده توسط مدل XGBoost با داده‌های تجربی واقعی مقایسه شد. نتایج نشان داد که انحرافات بین مقادیر پیش‌بینی‌شده و واقعی عمدتاً کمتر از ±۲ درصد است (شکل ۴). این موضوع بیانگر پایداری آماری و دقت قابل اعتماد مدل است، هرچند باید در تعمیم نتایج به فرمولاسیون‌های جدید با احتیاط عمل کرد؛ زیرا اندازهٔ مجموعه‌داده محدود بوده است.



شکل ۳. نمودار خلاصهٔ SHAP برای اهمیت ویژگی‌ها

[Downloaded from journal.nkums.ac.ir on 2026-11-23]



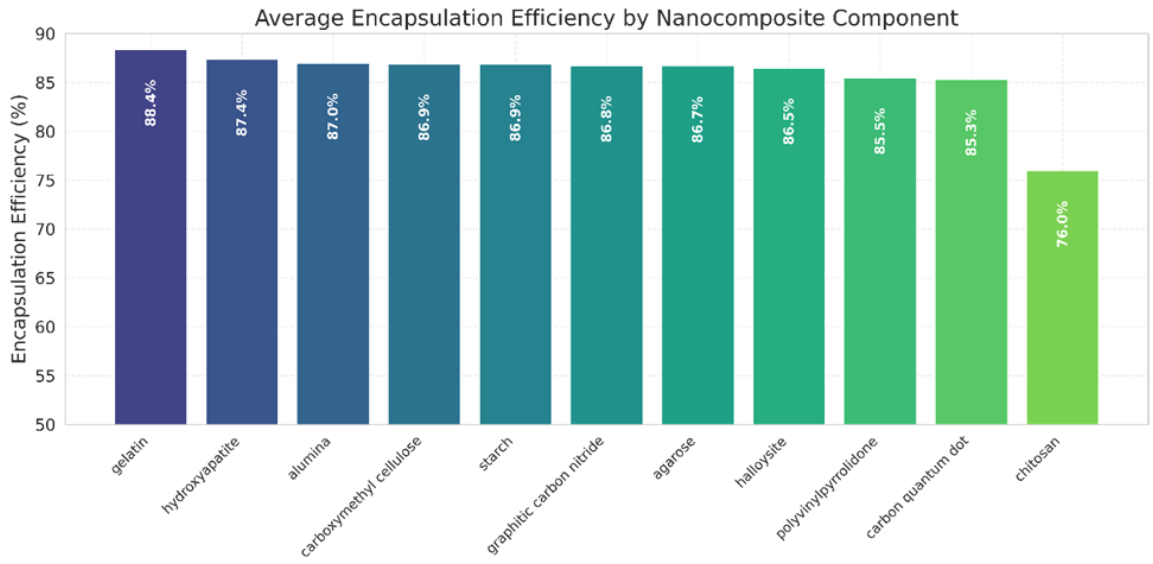
شکل ۴. درصد بازده پوشش‌دهی واقعی در مقابل درصد پیش‌بینی‌شده با معیارهای عملکرد

این نتایج تأیید می‌کند که استفاده از زیست‌پلیمرهای انعطاف‌پذیر به افزایش کارایی و ایمنی منجر می‌شود، درحالی‌که حامل‌های معدنی سخت معمولاً سمیت بیشتری دارند. مقایسهٔ اثر نسبی اجزا در شکل ۵ آورده شده است.

فراتر از این محدوده، بازده پوشش‌دهی به‌طور چشمگیری کاهش یافت. ذرات بزرگ‌تر از ۴۰۰ نانومتر یا حامل‌های با بار منفی، همواره کارایی ضعیفی نشان دادند. این یافته‌ها اصول طراحی سودمندی برای دستیابی به تعادل میان ایمنی و بازده در مهندسی نانوکامپوزیت‌ها ارائه می‌کنند [۴۳].

تحلیل سطح پاسخ سه‌بعدی (3D Response Surface Analysis) تعاملات غیرخطی میان اندازه ذره، پتانسیل زتا و EE% را آشکار کرد (شکل ۶). این تحلیل محدوده‌ای بهینه (Sweet Spot) را در اندازهٔ ۱۰۰ تا ۳۰۰ نانومتر و پتانسیل زتای مثبت ۲۰+ تا ۴۰+ میلی‌ولت تعیین کرد.

غربالگری مجازی: مدل‌های آموزش‌دیده برای غربالگری ۱۲۰ فرمولاسیون ترکیبی مجازی استفاده شدند. سه ترکیب برتر به‌عنوان طرح‌های بهینه پیشنهادی شناسایی شدند که در جدول ۳ نشان داده شده‌اند.



شکل ۵. مقایسهٔ راندمان اجزا

جدول ۳. سه فرمولاسیون برتر برای غربالگری مجازی از نظر LE%، EE% و سمیت سلولی

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

فرمولاسیون	LE% پیش‌بینی‌شده	EE% پیش‌بینی‌شده	ریسک سمیت سلولی
کیتوزان-g/PEG/ C_3N_4/هیدروکسی‌آپاتیت	۵۱/۲	۸۹/۷	پایین (احتمال = ۰/۰۹)
آگاروز/PEG/سیلیکای مزومتخلخل	۴۷/۸	۸۶/۳	پایین (احتمال = ۰/۱۲)
ژلاتین/کیتوزان Fe_3O_4/	۴۴/۹	۸۳/۱	متوسط (احتمال = ۰/۳۴)

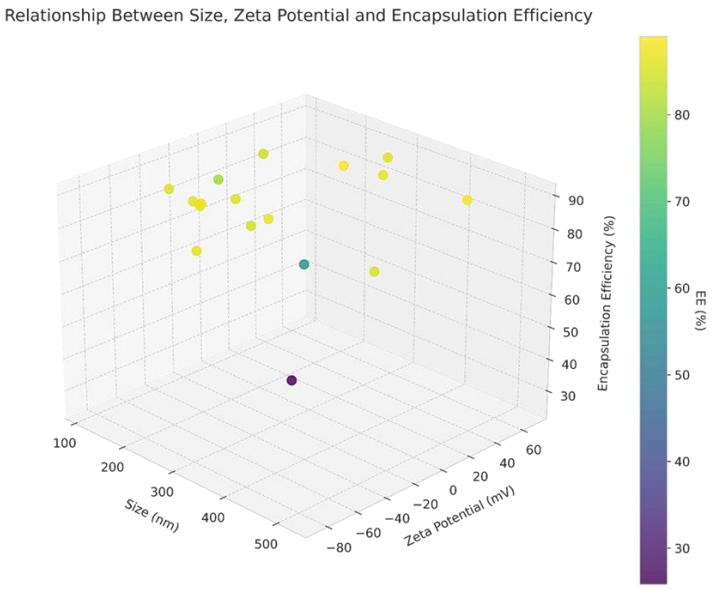
بحث

یافته‌های این پژوهش می‌تواند نیاز به آزمایش‌های تکراری و زمان‌بر را کاهش دهد، چرخهٔ طراحی نانوداروها را تسریع و مسیر توسعهٔ نانوداروهای مبتنی‌بر ۵-فلوئوروراوراسیل را به‌سوی پزشکی دقیق هدایت کند. این تحقیق بخشی از جریان روبه‌رشد استفاده از هوش محاسباتی در پژوهش‌های نانوتراپی و گذار به‌سمت انکولوژی دقیق به‌شمار می‌رود. مهم‌ترین یافته، برتری مدل‌های XGBoost و Random Forest در پیش‌بینی هم‌زمان LE%، EE% و سمیت سلولی است عملکرد مدل‌ها استثنایی بود؛ XGBoost با R^2=۰/۹۱ و MAE=۱/۸۲ برای LE% و R^2=۰/۸۹ با MAE=۲/۰۵ برای EE%، حدود ۹۰ درصد واریانس داده‌ها را تبیین کرد. Random Forest با F1=۰/۸۸ و AUC=۰/۹۲ تمایز دقیقی میان سامانه‌های سمی و غیرسمی فراهم آورد. این مقادیر نشان می‌دهد که مدل حدود ۹۰ درصد واریانس داده‌های تجربی را تبیین کرده است و اکثر خطاهای پیش‌بینی در بازهٔ ±۲ درصد قرار داشته‌اند. این مقادیر بیانگر حساسیت و ویژگی بالای مدل در تفکیک سامانه‌های سمی و غیرسمی هستند.

مدل‌های (XGBoost، Random Forest، LightGBM) Ensemble در مقایسه با SVR و ANN، عملکرد پایدارتری نشان دادند که در داده‌های کوچک دچار بیش‌برازش شدند. تغییرات میان folds ناچیز بود و قابلیت اطمینان پیش‌بینی‌ها را تقویت کرد. این نتایج با یافته‌های پیشین هم‌خوان است که رویکردهای تجمعی را برای داده‌های زیست‌پزشکی پیچیده مؤثرتر معرفی کرده‌اند. مقایسه گرافیکی عملکرد الگوریتم‌ها در شکل ۲ نمایش داده شده است.

تحلیل SHAP، پتانسیل زتا را به‌عنوان اثرگذارترین ویژگی شناسایی کرد؛ بارهای مثبت> ۳۰+ mV هم EE% را کاهش و هم سمیت را افزایش می‌دادند. این نتیجه با مطالعات پیشین مطابقت دارد که نشان داده‌اند ذرات با بار مثبت بالا غشای سلولی را تخریب می‌کنند . نمودار خلاصهٔ SHAP (شکل ۳) اهمیت نسبی هر عامل را رتبه‌بندی و تأیید می‌کند که پتانسیل زتا اثرگذارترین متغیر در هر دو شاخص پوشش‌دهی و سمیت است.

اندازه ذرات نیز کلیدی بود؛ محدوده ۱۰۰-۳۰۰ nm EE% را بهینه کرد، اما ذرات >۴۰۰ nm به‌دلیل تجمع، کارایی را مختل کردند. تحلیل سطح پاسخ سه‌بعدی (شکل ۶) تعاملات غیرخطی میان اندازه، زتا و EE% را آشکار و محدوده بهینه (۲۰+ تا ۴۰+mV ، ۱۰۰-۳۰۰ nm) را تعیین کرد. فراتر از این محدوده، بازده پوشش‌دهی به‌طور چشمگیری کاهش یافت.



شکل ۶. رابطهٔ سه‌بعدی اندازه، پتانسیل زتا و EE%

تفاوت عملکرد اجزا بینش‌های ساختاری ارائه داد؛ زیست‌پلیمرها (کیتوزان ۱/۵+ درصد)، PEG (۱/۹+ درصد)، سیلیکای مزومتخلخل (۲/۹+ درصد) از طریق پیوند هیدروژنی، پایداری فضایی و π-π stacking، EE% را بهبود بخشیدند. هیدروکسی‌آپاتیت (۷/۱+ درصد) و g-C_3N_4 (۱/۳+ درصد) بر LE% مؤثر بودند. نانولوله‌های هالویسیت (SHAP=+۰/۴۱) سمیت را افزایش دادند. مقایسهٔ بازدهی اجزا در شکل ۵ آمده است. این نتایج تأیید می‌کند که زیست‌پلیمرهای انعطاف‌پذیر کارایی و ایمنی را افزایش می‌دهند.

اعتبارسنجی با داده‌های تجربی، دقت را تأیید کرد. EE% ژلاتین PEG/g-C_3N_4 (پیش‌بینی ۸۹/۲ درصد ≈ تجربی ۸۹/۰±۱/۲ درصد) مطابقت عالی داشت، اما اختلاف کیتوزان/فولیک اسید (۶۸/۳ درصد نسبت به ۵۹ درصد) فقدان توصیفگرهای pH را نشان داد. انحرافات

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

±۲ درصد عمدتاً، پایداری آماری مدل را اثبات می‌کند (شکل ۴). غربالگری ۱۲۰ فرمولاسیون مجازی، کیتوزان/C_3N_4-g/PEG/هیدروکسی‌آپاتیت را با EE%= ۸۹/۷ و ریسک سمیت ۰/۰۹ پیشنهاد داد.

این یافته‌ها نقش ML در طراحی پیش‌بینانهٔ نانوداروها را برجسته می‌کند. بیوپلیمرها (کیتوزان/PEG)، اندازهٔ ۱۰۰-۳۰۰ nm، پرهیز از هالوسیت/بار بالا، با وجود n= ۲۰ و هم‌بستگی ذاتی (r=۰/۹۳)، پتانسیل ML مشهود است.

نتیجه‌گیری

نتایج این پژوهش نشان داد که مدل XGBoost با دستیابی به بالاترین دقت پیش‌بینی (R^2=۰/۹۱ برای LE% و R^2 =۰/۸۹ برای EE%) و Random Forest با عملکرد برتر در طبقه‌بندی سمیت سلولی (F1=۰/۸۸، AUC=۰/۹۲) به‌عنوان مدل‌های بهینه برای پیش‌بینی ویژگی‌های نانوکامپوزیت‌های حامل ۵-فلوئورواوراسیل شناسایی شدند. تحلیل SHAP اصول طراحی کلیدی را آشکار کرد. (i): اندازهٔ ذرات در بازهٔ ۱۰۰-۳۰۰ نانومتر، (ii): پتانسیل زتای مثبت ۲۰+ تا ۴۰+ میلی‌ولت، (iii) استفاده از زیست‌پلیمرهای کیتوزان-g/PEG/C_3N_4/هیدروکسی‌آپاتیت به‌عنوان ترکیب برتر با EE% پیش‌بینی‌شده ۸۹/۷ درصد و ریسک سمیت کم (احتمال = ۰/۰۹). غربالگری مجازی ۱۲۰ فرمولاسیون فرضی نیز فرمولاسیون کیتوزان/C_3N_4-g/PEG/هیدروکسی‌آپاتیت را به‌عنوان کاندیدای برتر برای سنتز و آزمون‌های in vitro/in vivo پیشنهاد داد.

این چهارچوب یادگیری ماشین چندهدفه نه‌تنها دقت پیش‌بینی LE%، EE% و سمیت را به‌طور هم‌زمان بهبود بخشید، بلکه با استخراج روابط مکانیسمی میان ویژگی‌های فیزیکوشیمیایی و عملکرد درمانی، مسیر طراحی منطقی و سریع نانوحامل‌های ضدسرطان را هموار کرد. یافته‌های حاضر ظرفیت زیاد رویکردهای محاسباتی در کاهش آزمایش‌های تجربی زمان‌بر و هدایت بهینه‌سازی نانوداروها به‌سوی پزشکی دقیق را تأیید می‌کند.

این مدل‌ها توانستند ضمن افزایش دقت پیش‌بینی، بینش‌های مکانیسمی ارزشمندی دربارهٔ رابطه بین ویژگی‌های فیزیکوشیمیایی و رفتار درمانی نانوکامپوزیت‌ها ارائه دهند. باوجود این دستاوردها، مطالعهٔ حاضر با چند محدودیت اساسی روبه‌رو است: اندازهٔ کوچک مجموعه‌داده (n=۲۰) موجب کاهش قابلیت تعمیم مدل می‌شود و نبود برخی توصیفگرهای کلیدی مانند نسبت دارو به پلیمر، وزن مولکولی، حساسیت به pH و شرایط سنتز، دقت مدل را محدود کرده است. افزون‌بر این، هم‌بستگی ذاتی میان پتانسیل زتا و اندازهٔ ذره، تفسیر نتایج را دشوار کرد و نبود اعتبارسنجی آزمایشگاهی گسترده سبب می‌شود که پیش‌بینی‌ها نیازمند تأیید تجربی در مقیاس وسیع‌تر باشند.

چشم‌اندازهای آینده شامل گسترش مجموعه‌داده‌ها با استفاده از سامانه‌های متنوع نانوکامپوزیتی و تکرارهای آزمایشی بیشتر است. همچنین استفاده از توصیفگرهای جدید مانند نرخ آزادسازی دارو، ویژگی‌های شیمیایی پلیمر و وزن مولکولی می‌تواند دقت پیش‌بینی را افزایش دهد. ترکیب یادگیری ماشین با شبیه‌سازی‌های دینامیک مولکولی و مدل‌سازی QSAR امکان توسعهٔ چهارچوب‌های پیش‌بینی ترکیبی را فراهم می‌کند و بهره‌گیری از یادگیری انتقالی می‌تواند با استفاده از داده‌های مشابه در حوزهٔ نانوداروها، کمبود داده را جبران کند. علاوه‌بر این، اعتبارسنجی آزمایشگاهی مبتنی‌بر حلقهٔ بازخوردی، امکان بهینه‌سازی تدریجی و مستمر مدل را ایجاد خواهد کرد.

در مجموع، رویکرد یکپارچه‌ی ارائه‌شده می‌تواند فرایند توسعهٔ دارو را از روش‌های سنتی مبتنی‌بر آزمون و خطا به‌سمت طراحی نظام‌مند و پیش‌بینانه در نانوداروهای هدفمند سوق دهد و گامی مؤثر در تحقق پزشکی دقیق (Precision Oncology) به شمار آید.

سپاسگزاری

نویسندگان از دانشگاه زابل برای حمایت‌های ارائه‌شده در انجام این پژوهش صمیمانه سپاسگزارند.

ملاحظات اخلاقی

این مطالعه بر اساس منابع منتشرشده تحلیل و داده پردازی شده است و نویسندگان متعهد به رعایت استانداردهای اخلاقی نشر علمی، از جمله شفافیت و عدم تعارض منافع، می‌باشند.

تعارض منافع

نویسندگان هیچ‌گونه تعارض منافعی با یکدیگر ندارند.

مشارکت نویسندگان

عباس رهدار: برنامه‌نویسی و اجرای تحلیل‌ها با پایتون، تفسیر آماری نتایج و مدیریت پروژه را عهده‌دار بود؛ مریم شیرزاد: تحلیل داده‌ها و بازبینی انتقادی و نهایی را انجام داد؛ مهراب پور مددی: مفهوم‌سازی مطالعه و نگارش پیش‌نویس اولیه مقاله را بر عهده داشت؛ تمامی نویسندگان نسخه نهایی را بررسی و تأیید کردند.

حمایت مالی

این پژوهش از حمایت مالی خارجی بهره‌مند نبوده است.

References

1. Bertão AR, Teixeira F, Ivasiv V, Parpot P, Almeida-Aguiar C, Fonseca AM, et al. Machine learning-assisted optimization of drug combinations in zeolite-based delivery systems for melanoma therapy. ACS Appl Mater Interfaces. 2024;16(5):5696-707. [DOI: 10.1021/acsami.3c18224] [PMID: 38271191]
2. Hajareh Haghighi F, Binaymotlagh R, Fratoddi I, Chronopoulou L, Palocci C. Peptide-hydrogel nanocomposites for anti-cancer drug delivery. Gels. 2023;9(12):953. [DOI: 10.3390/gels9120953] [PMID: 38131939]
3. Hristova-Panusheva K, Xenodochidis C, Georgieva M, Krasteva N. Nanoparticle-mediated drug delivery systems for precision targeting in oncology. Pharmaceuticals. 2024;17(6):677. [DOI: 10.3390/ph17060677] [PMID: 38931344]
4. Abed A, Karimi M, Nejati M, Hamblin MR, Mirzaei SA, Sarvizadeh M, et al. Synthesis of 5-fluorouracil (5-FU) coated platinum nanoparticles and apoptotic effects on U87 human glioblastoma cells. Cancer Cell Int. 2025;25(1):280. [DOI: 10.1186/s12935-025-03893-w] [PMID: 40713567]

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

5. Izadiyan Z, Webster TJ, Kia P, Kalantari K, Misran M, Rasouli E, et al. Nanoemulsions based therapeutic strategies: enhancing targeted drug delivery against breast cancer cells. Int J Nanomedicine. 2025;20:6133-62. [DOI: 10.2147/ijn.s488545] [PMID: 40385492]
6. Noruzpuor M, Asghari Zakaria R, Zare N, Bourang S, Ebrahimi HA, Granados-Principal S. Delivery of Moringa oleifera extract via PLA-PEG-FA/chitosan-PLA NPs into breast cancer cell lines. BioNanoScience. 2025;15(2):287. [DOI: 10.2139/ssrn.5036589]
7. Sun L, Liu H, Ye Y, Lei Y, Islam R, Tan S, et al. Smart nanoparticles for cancer therapy. Signal Transduct Target Ther. 2023;8(1):418. [DOI: 10.1038/s41392-023-01642-x] [PMID: 37919282]
8. Li B, Shao H, Gao L, Li H, Sheng H, Zhu L. Nano-drug co-delivery system of natural active ingredients and chemotherapy drugs for cancer treatment: a review. Drug Deliv. 2022;29(1):2130-61. [DOI: 10.1080/10717544.2022.2094498] [PMID: 35815678]
9. Păduraru L, Panainte AD, Peptu CA, Apostu M, Vieriu M, Bibire T, et al. Smart drug delivery systems based on cyclodextrins and chitosan for cancer therapy. Pharmaceuticals. 2025;18(4):564. [DOI: 10.3390/ph18040564] [PMID: 40283999]
10. Chen J, Han H, Li L, Chen Z, Liu X, Li T, et al. Prediction of cancer cell line-specific synergistic drug combinations based on multi-omics data. PeerJ. 2025;13:e19078. [DOI: 10.7717/peerj.19078] [PMID: 40028209]
11. Khodadadi F, Taghizadeh F, Hashemi Baghi A, Ayyoubzadeh SM, Dadashzadeh S, Haeri A,et al. Leveraging ensemble machine learning models (XGBoost and random forest) and genetic algorithms to predict factors contributing to the liposomal entrapment of therapeutics. Nanoscale. 2025;17(38):22271-90. [DOI: 10.1039/D5NR01741F]
12. Muratov EN, Bajorath J, Sheridan RP, Tetko IV, Filimonov D, Poroikov V, et al. QSAR without borders. Chem Soc Rev. 2020;49(11):3525-64. [DOI: 10.1039/d0cs00098a] [PMID: 32356548]
13. Wang H, Di W, Gao X, Guo Y, Tang T, Bai X, et al. Materials, syntheses and biomedical applications of nano-quercetin formulations: a comprehensive literature review. Int J Nanomedicine. 2025;20:8729-64. [DOI: 10.2147/ijn.s517079] [PMID: 40635717]
14. Yildirim E, Cakir I, Ileri-Ercan N. Machine learning analysis of cytotoxicity determinants in nanoparticle-based rheumatoid arthritis therapies. Mol Pharm. 2025;22(11):6703-13. [DOI: 10.1021/acs.molpharmaceut.5c00661] [PMID: 41129739]
15. Alqarni S, Huwaimel B. Predicting PLGA nanoparticle size and zeta potential in synthesis for application of drug delivery via machine learning analysis. Sci Rep. 2025;15(1):20765. [DOI: 10.1038/s41598-025-06872-3] [PMID: 40594444]
16. Park J, Rashid S, Copsey H, Tran L, Zabeo A, Hristozov D, et al. NanoToxRadar: a multitarget nano-QSAR model for predicting the cytotoxicity values of multicomponent nanoparticles. ACS Nanoscience Au. 2025;5(5):344-52. [DOI: 10.1021/acsnanoscienceau.5c00035] [PMID: 41112870]
17. Bereyhi M, Zare-Dorabei R, Safarifard V. A biocompatible NMOF-based drug delivery system designed for the sustained and controlled release via pH-responsive poorly soluble drug 5-fluorouracil. Langmuir. 2025;41(10):6544-52. [PMID: 10.1021/acs.langmuir.4c03953] [PMID: 40056115]
18. Pourmadadi M, Ajalli N, Nouri M, Shirazi M. Enhanced drug delivery of 5-fluorouracil using a GO-PVP-SA nanocomposite for targeted colorectal cancer treatment. BioNanoScience. 2025;15(1):194. [DOI: 10.1007/s12668-025-01800-1]
19. Pourmadadi M, Poorkhalili P, Sorourian M, Sorourian G, Ghaderi R, Mehrabi MG, Ajalli N. The smart nanocarrier containing zein/starch co-biopolymers enhanced by graphitic carbon nitride; exploring opportunities in brain cancer treatment. Int J Biol Macromol. 2024;274:133275. [DOI: 10.1016/j.ijbiomac.2024.133275] [PMID: 38906350]
20. Pourmadadi M, Garousi NA, Abdouss M, Rahdar A, Fathi-Karkan S, Pandey S. PEG-modified Fe2O3 coated agarose hydrogel: a synthesized nanocomposite for regulated 5-fluorouracil delivery. Int J Biol Macromol. 2024;276:133900. [DOI: 10.1016/j.ijbiomac.2024.133900] [PMID: 39019377]
21. Zavareh HS, Pourmadadi M, Moradi A, Yazdian F, Omidi M. Chitosan/carbon quantum dot/aptamer complex as a potential anticancer drug delivery system towards the release of 5-fluorouracil. Int J Biol Macromol. 2020;165:1422-30. [DOI: 10.1016/j.ijbiomac.2020.09.166] [PMID: 32987067]
22. Ahmari A, Pourmadadi M, Yazdian F, Rashedi H, Khanbeigi KA. A green approach for preparation of chitosan/hydroxyapatite/graphitic carbon nitride hydrogel nanocomposite for improved 5-FU delivery. Int J Biol Macromol. 2024;258:128736. [DOI: 10.1016/j.ijbiomac.2023.128736] [PMID: 38101677]
23. Shamsabadipour A, Pourmadadi M, Rashedi H, Yazdian F, Navaei-Nigjeh M. Nanoemulsion carriers of porous γ-alumina modified by polyvinylpyrrolidone and carboxymethyl cellulose for pH-sensitive delivery of 5-fluorouracil. Int J Biol Macromol. 2023;233:123621. [DOI: 10.1016/j.ijbiomac.2023.123621] [PMID: 36773864]
24. Ghasemizadeh H, Pourmadadi M, Yazdian F, Rashedi H, Navaei-Nigjeh M, Rahdar A, et al. Novel carboxymethyl cellulose-halloysite-polyethylene glycol nanocomposite for improved 5-FU delivery. Int J Biol Macromol. 2023;232:123437. [DOI: 10.1016/j.ijbiomac.2023.123437] [PMID: 36708898]
25. Rahmani M, Pourmadadi M, Abdous M, Radhar A, Díez Pascual AM. Gelatin/polyethylene glycol/g-C3N4 hydrogel with olive oil as a sustainable and biocompatible nanovehicle for targeted delivery of 5-fluorouracil. Ind Crops Prod. 2024;208:117912. [DOI: 10.1016/j.indcrop.2023.117912]
26. Ghotbi M, Pourmadadi M, Yazdian F, Hallajsani A. Fabrication and characterization of starch/agarose biopolymers containing graphene oxide towards the release of 5-fluorouracil in cancer treatment. Inorg Chem Commun. 2024;170:113119. [DOI: 10.1016/j.inoche.2024.113119]
27. Khoddam Z, Pourmadadi M, Abdouss M, Mazinani S, Behzadmehr R, Rahdar A, Aboudzadeh MA. Hybrid nanocarriers based on polyacrylic acid, polyvinyl pyrrolidone, and molybdenum disulfide for enhanced 5-fluorouracil delivery in lung cancer therapy. Inorg Chem Commun. 2024;159:111749. [DOI: 10.1016/j.inoche.2023.111749]
28. Ajalli N, Pourmadadi M, Yazdian F, Rashedi H, Navaei-Nigjeh M, Díez-Pascual AM. Chitosan/gamma-alumina/Fe3O4@5-FU nanostructures as promising nanocarriers: physiochemical characterization and toxicity activity. Molecules. 2022;27(17):5369. [DOI: 10.3390/molecules27175369]
29. Bayat F, Pourmadadi M, Eshaghi MM, Yazdian F, Rashedi H. Improving release profile and anticancer activity of 5-fluorouracil for breast cancer therapy using a double drug delivery system: chitosan/agarose/γ-alumina nanocomposite@double emulsion. J Cluster Sci. 2023;34(5):2565-77. [DOI: 10.1007/s10876-023-02405-y]
30. Mousavi M, Pourmadadi M, Abdouss M, Rahdar A, Fathi Karkan S, Pandey S. Gelatin/CMC/HAP nanocomposites based on double micro-emulsion for delivery of 5-FU: synthesis and chemical-physical characterization. BioNanoScience. 2024;14(5):5513-26. [DOI: 10.1007/s12668-024-01515-9]
31. Masnavi M, Pourmadadi M, Abdouss M, Rahdar A, Fathi

[Downloaded from journal.nkums.ac.ir on 2026-11-23]

Karkan S, Pandey S. Synthesis of nanocarrier based on chitosan/agarose biopolymers containing carbon quantum dot doped titanium dioxide for targeted delivery of 5-fluorouracil for brain cancer treatment. BioNanoScience. 2024;14(3):2264-74. [DOI: 10.1007/s12668-024-01593-9]
32. Abbaszadeh Garousi N, Pourmadadi M, Abdouss M, Rahdar A, Aboudzadeh MA. Hybrid hydrogels based on agarose, halloysite nanotubes and polyvinyl alcohol: an excellent nanocarrier for controlled release of 5-fluorouracil. Hybrid Adv. 2024;5:100139. [DOI: 10.1016/j.hybadv.2023.100139]
33. Ullah S, Azad AK, Nawaz A, Shah KU, Iqbal M, Albadrani GM, et al. 5-fluorouracil-loaded folic-acid-fabricated chitosan nanoparticles for site-targeted drug delivery cargo. Polymers. 2022;14(10):2010. [DOI: 10.3390/polym14102010]
34. Alhowyan AA, Kalam MA, Iqbal M, Raish M, El-Toni AM, Alkholief M, et al. Mesoporous silica nanoparticles coated with carboxymethyl chitosan for 5-fluorouracil ocular delivery: characterization, in vitro and in vivo studies. Molecules. 2023;28(3):1260. [DOI: 10.3390/molecules28031260] [PMID: 36770926]
35. Khan TA, Azad AK, Fuloria S, Nawaz A, Subramaniyan V, Akhlaq M, et al. Chitosan-coated 5-fluorouracil incorporated emulsions as transdermal drug delivery matrices. Polymers. 2021;13(19):3345. [DOI: 10.3390/polym13193345] [PMID: 34641162]
36. Yusefi M, Chan HY, Teow SY, Kia P, Lee-Kiun Soon M, Sidik NABC, et al. 5-fluorouracil encapsulated chitosan-cellulose fiber bionanocomposites: synthesis, characterization and in vitro analysis towards colorectal cancer cells. Nanomaterials. 2021;11(7):1691. [DOI: 10.3390/nano11071691] [PMID: 34203241]
37. Hudiță A, Radu IC, Zaharia C, Ion AC, Ginghină O, Gălățeanu B, et al. Bio- and hemo-compatible silk fibroin PEGylated nanocarriers for 5-fluorouracil chemotherapy in colorectal cancer: in vitro studies. Pharmaceutics. 2021;13(5):755. [DOI: 10.3390/pharmaceutics13050755] [PMID: 34069731]
38. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: with applications in R. New York: Springer; 2013. [Link]
39. Géron A. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. Sebastopol (CA): O'Reilly Media; 2022. [Link]
40. Honary S, Zahir F. Effect of zeta potential on the properties of nano-drug delivery systems: a review (Part 1). Trop J Pharm Res. 2013;12(2):255-64. [DOI: 10.4314/tjpr.v12i2.19]
41. Honary S, Zahir F. Effect of zeta potential on the properties of nano-drug delivery systems: a review (Part 2). Trop J Pharm Res. 2013;12(2):265-73. [DOI: 10.4314/tjpr.v12i2.20]
42. Tiwari G, Tiwari R, Sriwastawa B, Bhati L, Pandey S, Pandey P, et al. Drug delivery systems: an updated review. Int J Pharm Investig. 2012;2(1):2-11. [DOI: 10.4103/2230-973x.96920] [PMID: 23071954]
43. Myers RH, Montgomery DC. Response surface methodology. New York: Taylor & Francis; 1995. [Link]

[Downloaded from journal.nkums.ac.ir on 2026-11-23]