UNIVERSITY OF TORONTO
LESLIE DAN FACULTY OF PHARMACY

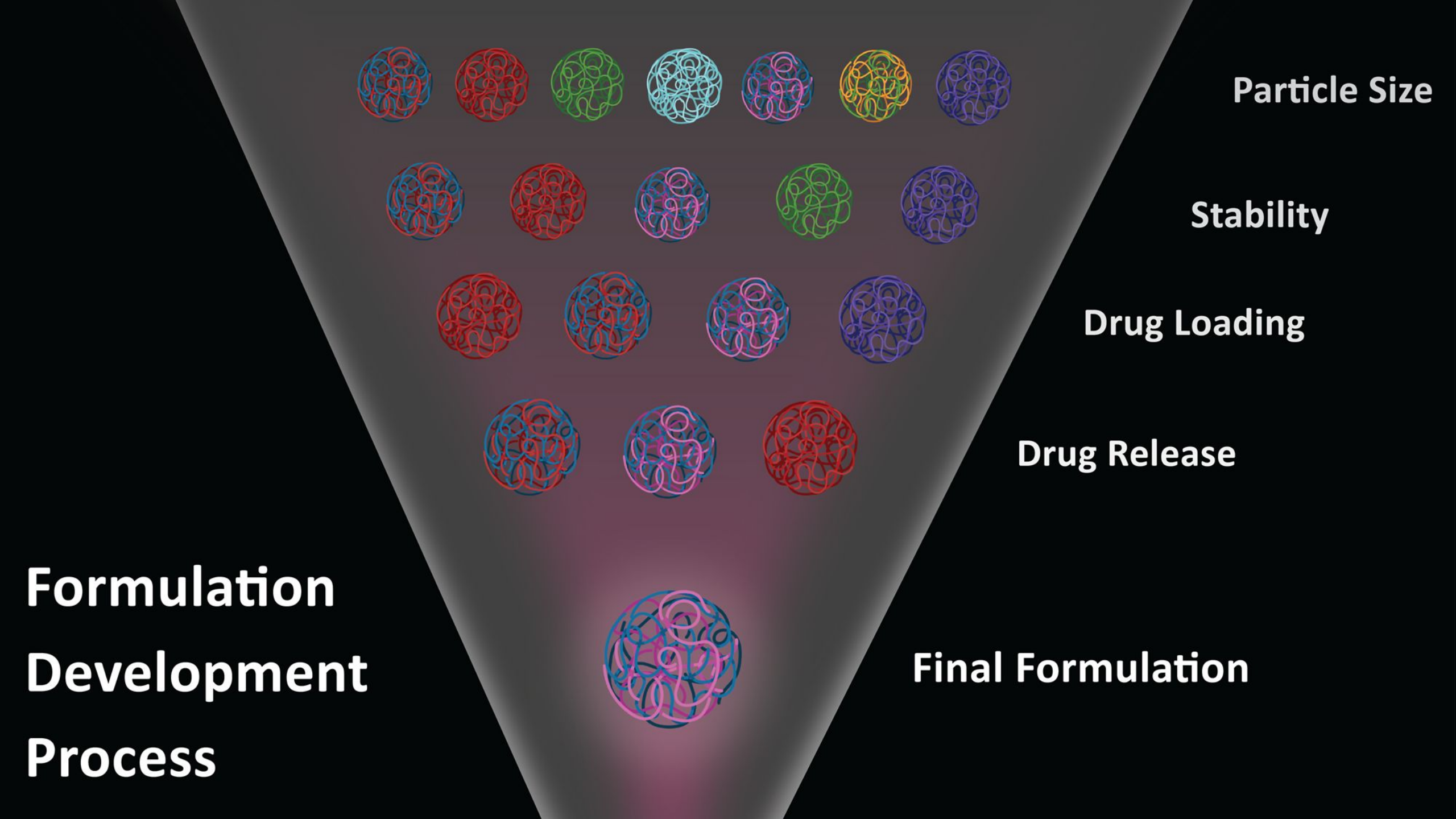
ALLEN LAB

Machine Learning Directed Drug Formulation Development

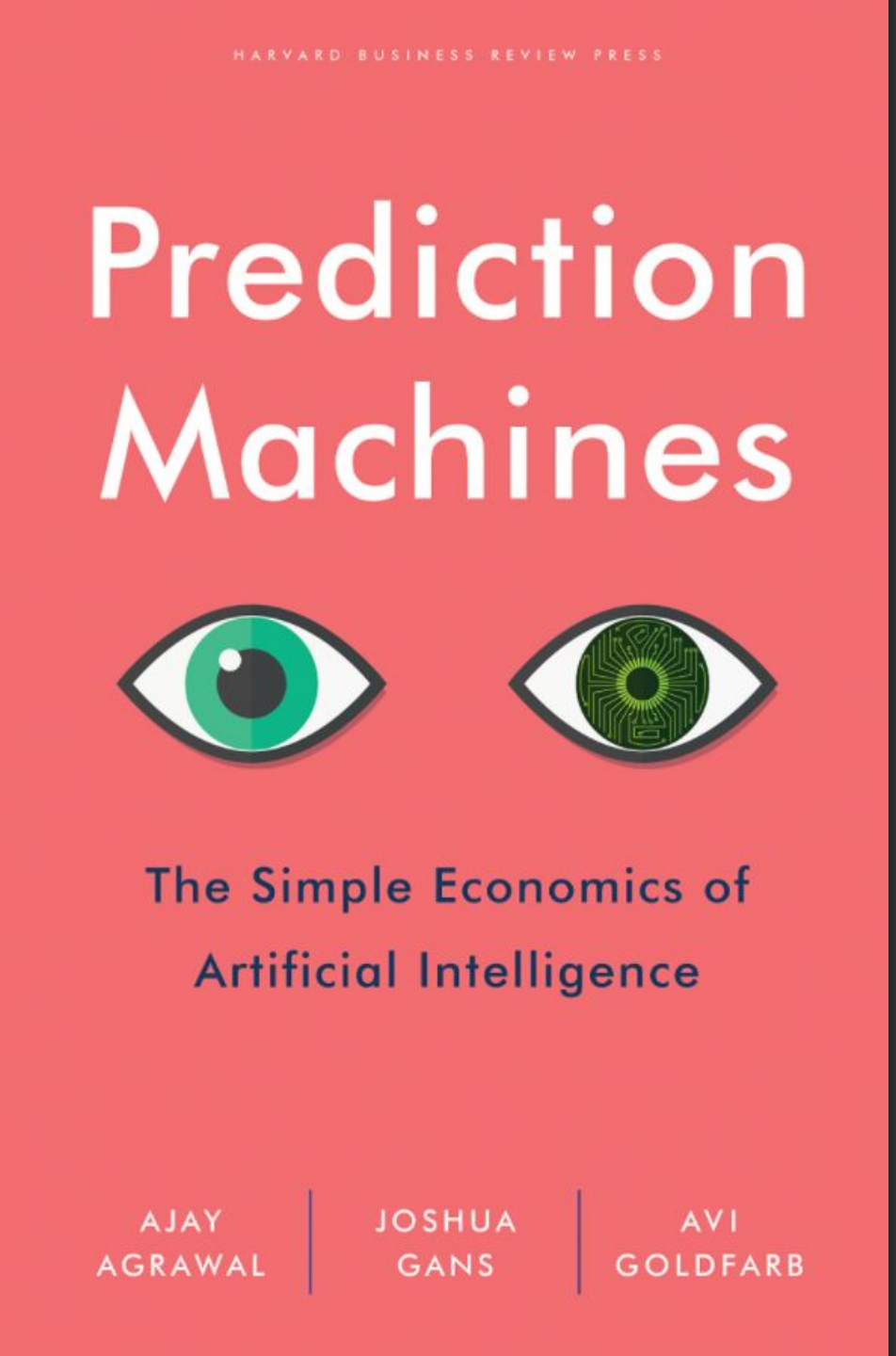
Zeqing Bao
Leslie Dan Faculty of Pharmacy
University of Toronto

the matter lab

July 2022


Particle Size
Stability
Drug Loading
Drug Release
Final Formulation
Formulation Development Process

Formulation Development



> Artificial intelligence is a technology of low-cost prediction and discovery

What prediction problems do we have in pharmaceutical sciences ?

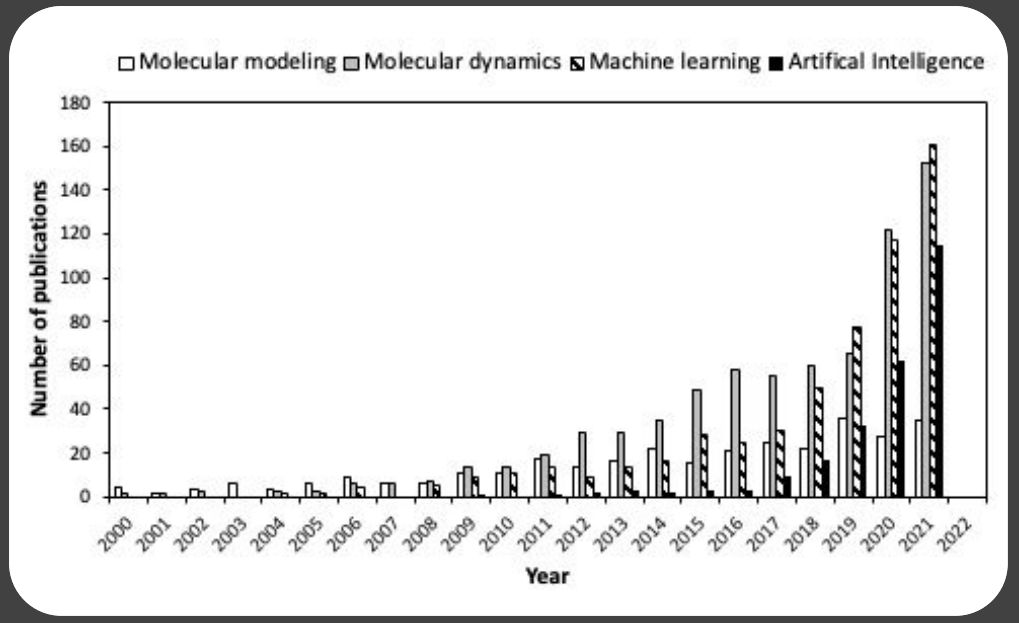
AI in Pharmaceutical Sciences

Journal of Controlled Release
Volumes 311–312, October 2019, Pages 326-327

Cover Story

Does artificial intelligence have the potential to transform drug formulation development?

Christine Allen, Pauric Bannigan



VECTOR INSTITUTE Roche

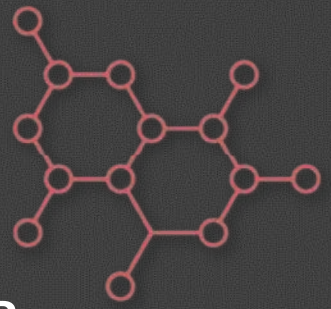
Creating an Intelligent and Connected Digital Workforce

Google | Pfizer

BenevolentAI | AstraZeneca

Identifying New Treatments for Chronic Kidney Disease & Idiopathic Pulmonary Fibrosis

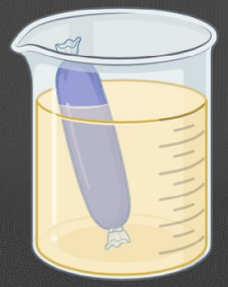
What would a ML approach look like?



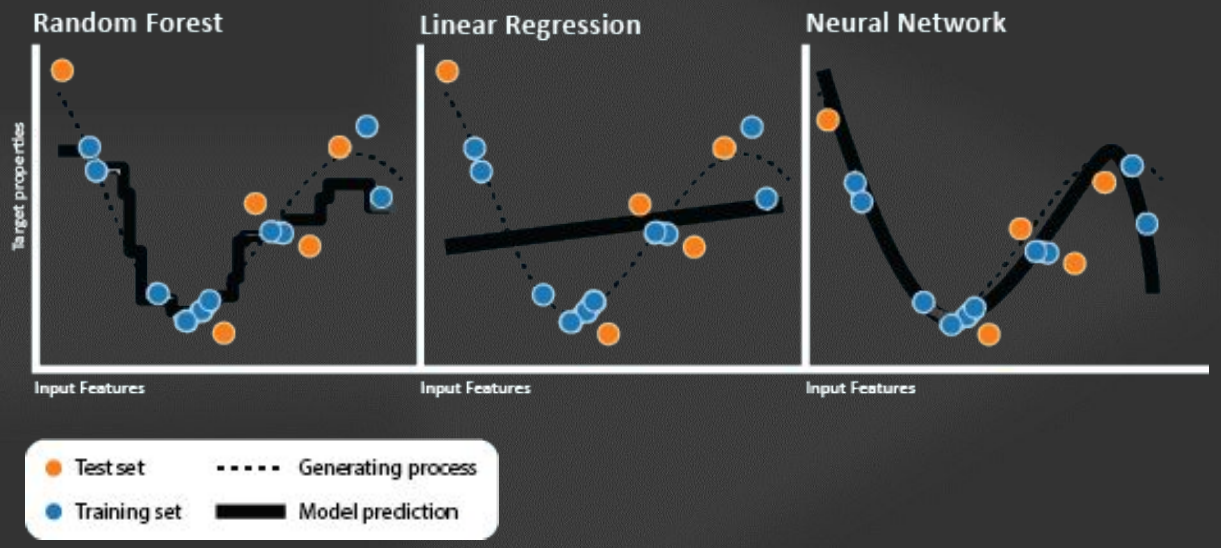
XLogP

Mw

Features



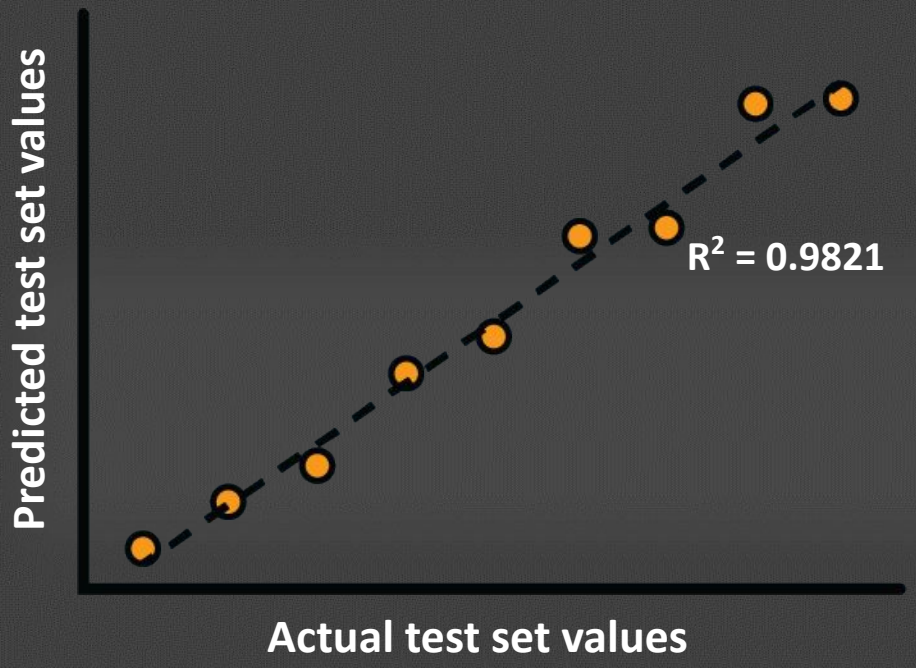
Release media composition

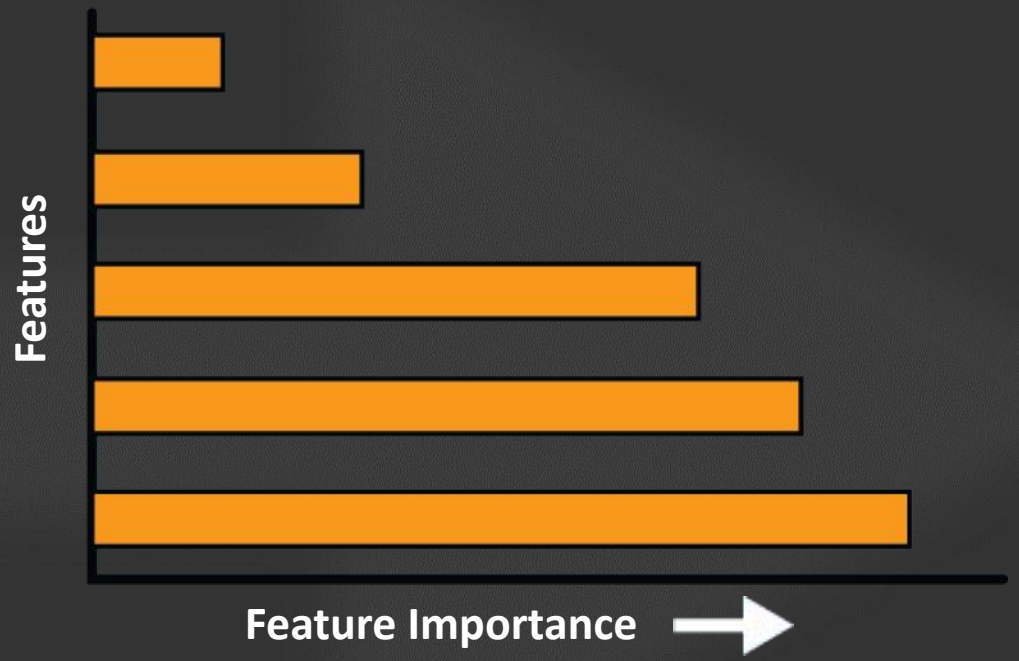


Dataset Construction

Model Selection

What would a ML approach look like?





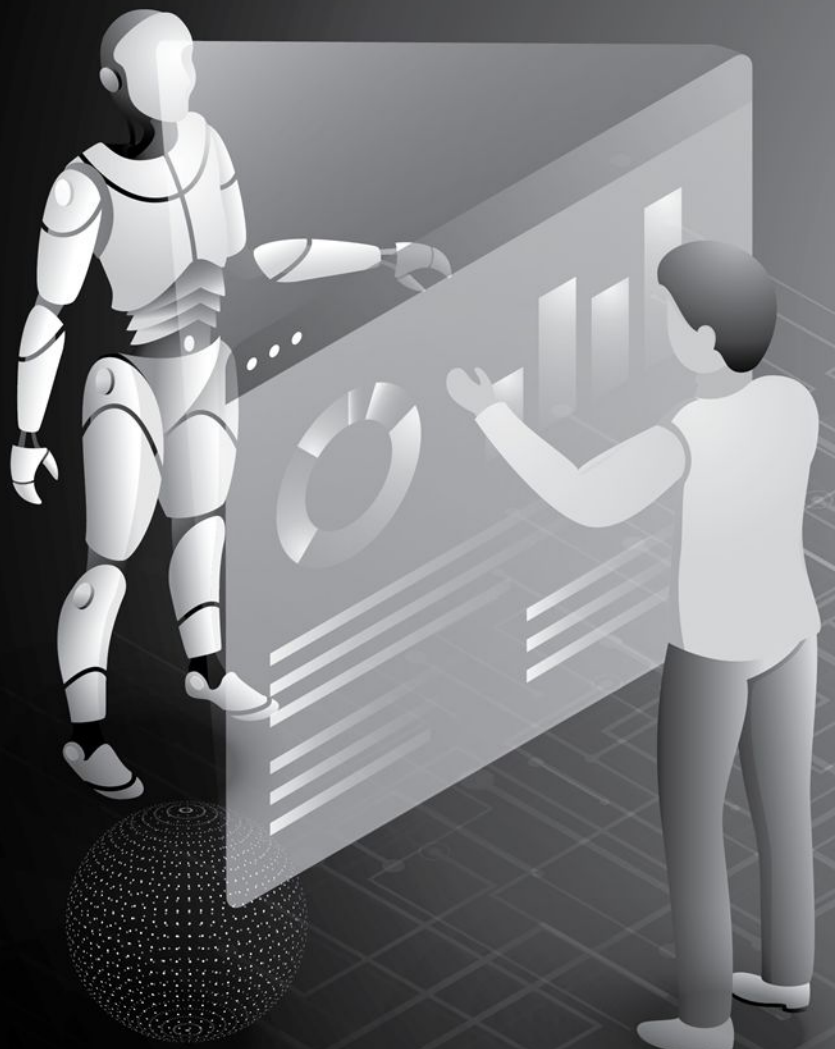
Model Evaluation

Model Interpretation

the matter lab

+

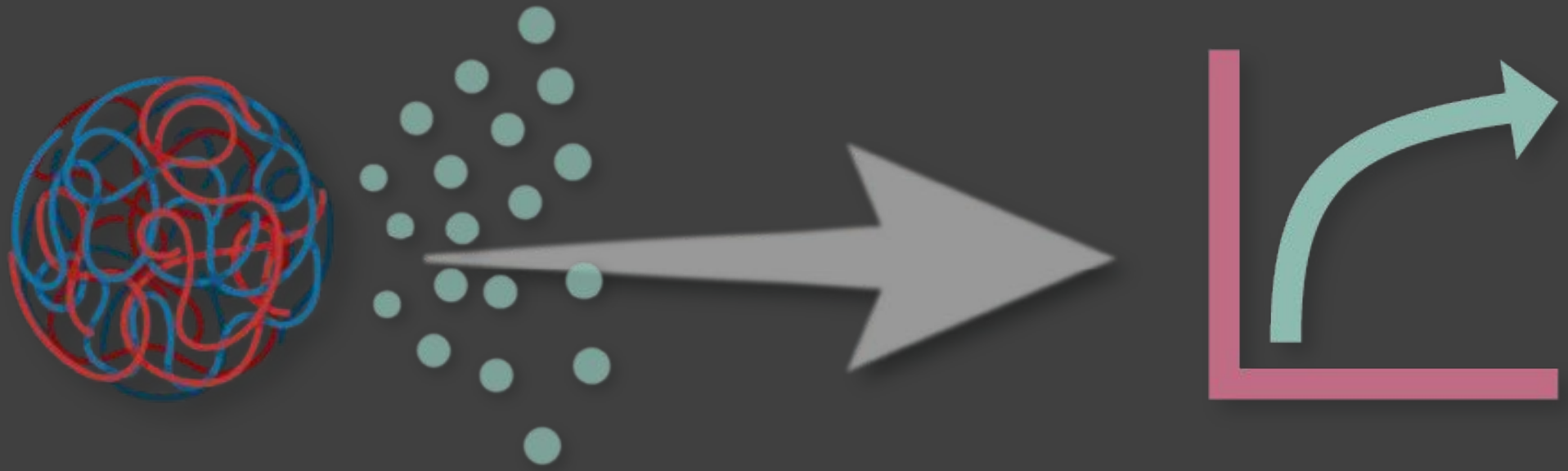
ALLEN LAB



Prediction of experimental conditions

Prediction of formulation performance

Prediction of Drug Release from MPs



Journal of Pharmaceutical Sciences
Volume 110, Issue 7, July 2021, Pages 2771-2777

Pharmaceutics, Drug Delivery and Pharmaceutical Technology

Poly(δ-valerolactone-*co*-allyl-δ-valerolactone) cross-linked microparticles: Formulation, characterization and biocompatibility

Zeqing Bao [a, 1], Sungmin Jung [a, 1], Jack Bufton [a], James C. Evans [a], Dean J. Aguiar [b], Christine Allen [a]

European Journal of Pharmaceutical Sciences
Volume 162, 1 July 2021, 105808

Cross-linked valerolactone copolymer implants with tailorable biodegradation, loading and *in vitro* release of paclitaxel

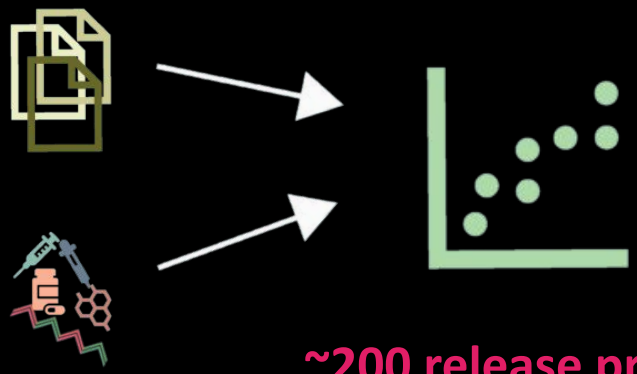
Jack Bufton [a, 2], Sungmin Jung [a, 1], James C. Evans [a], Zeqing Bao [a], Dean Aguiar [b], Christine Allen [a]

Data Collection

Clarivate Analytics

WEB OF SCIENCE™

Data extracted from 40 papers



~200 release profiles

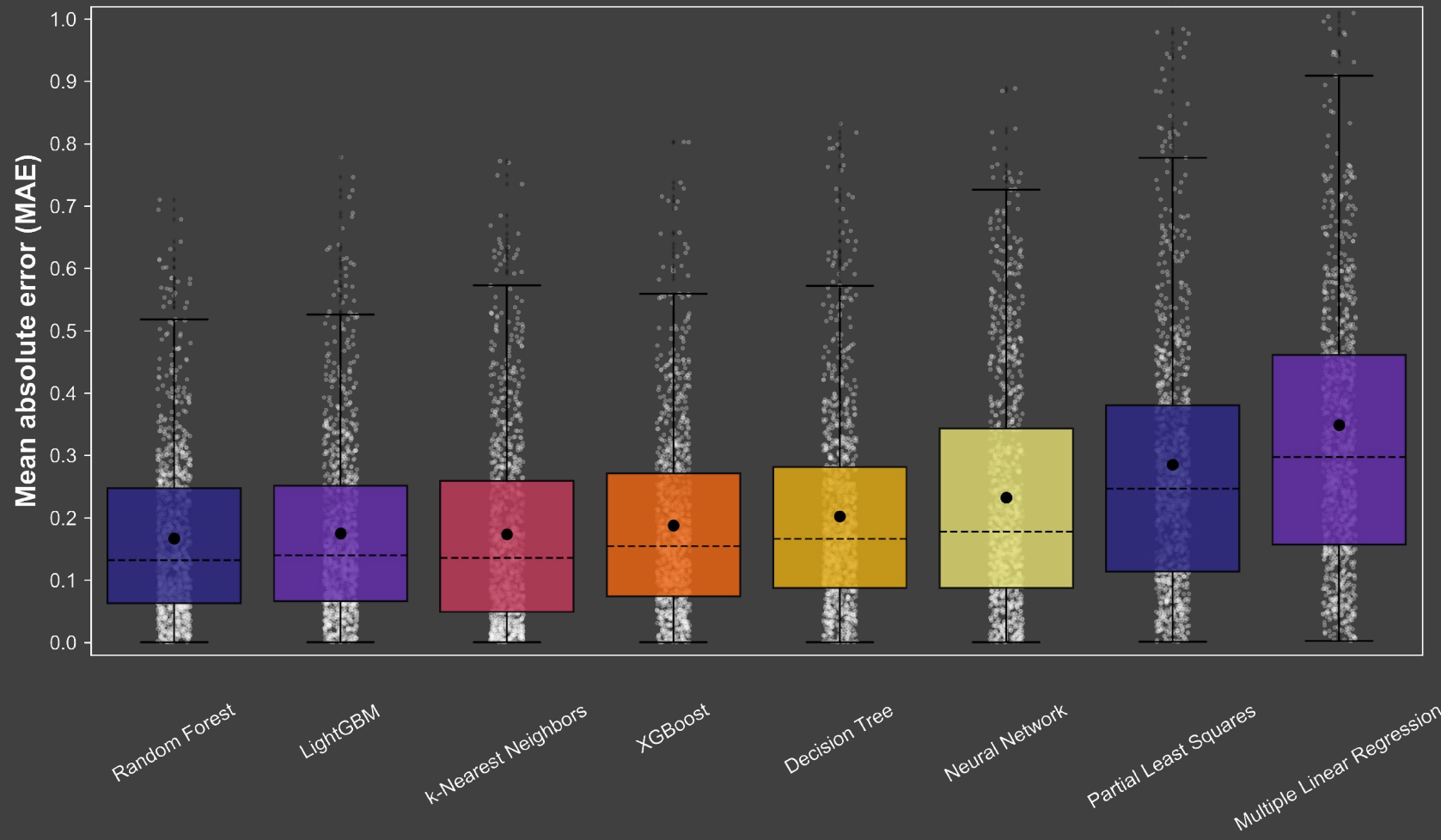
Training data:

~ 100 drug release

Validation data:

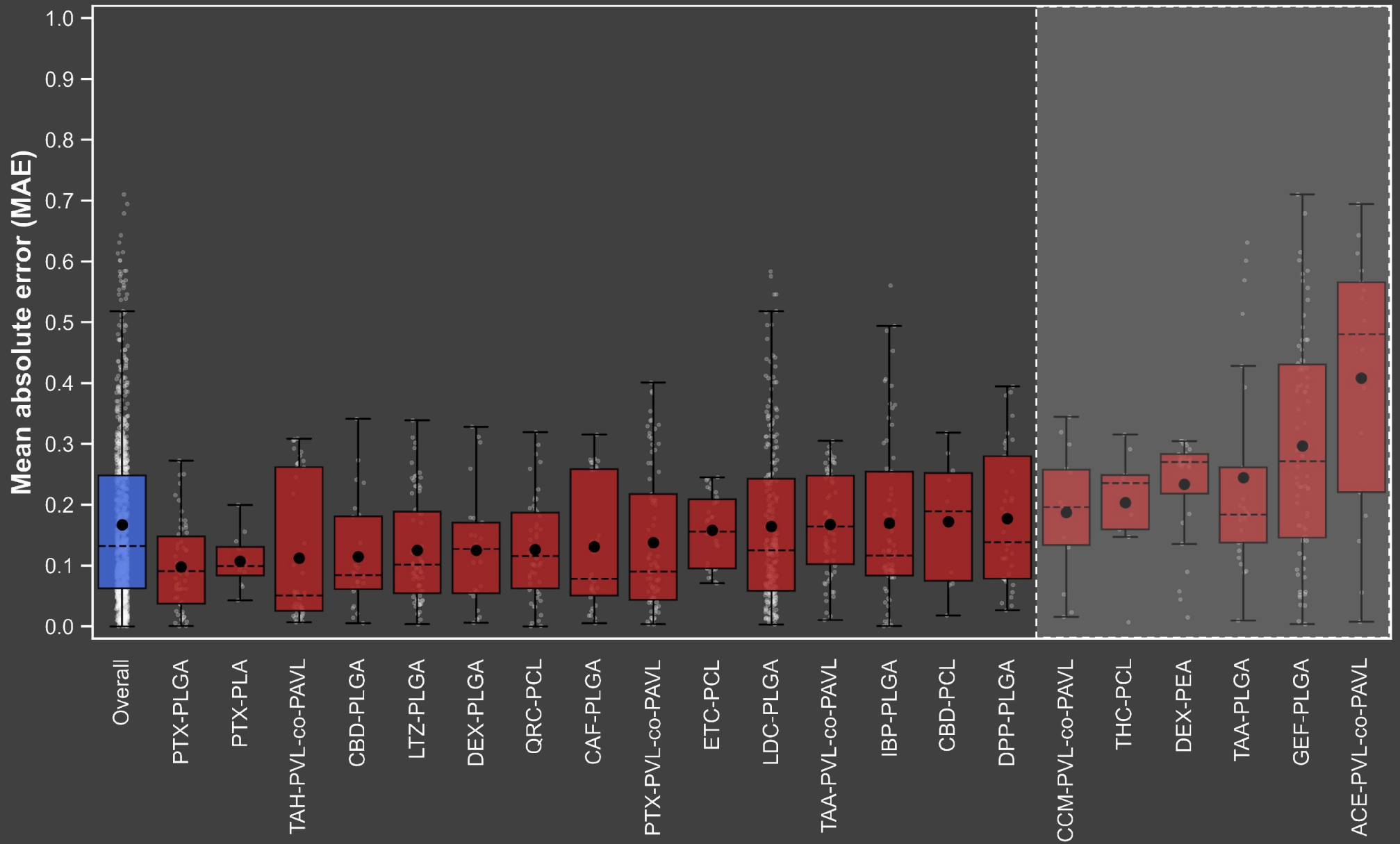
~ 80 drug release

ML Model Screening and Selection

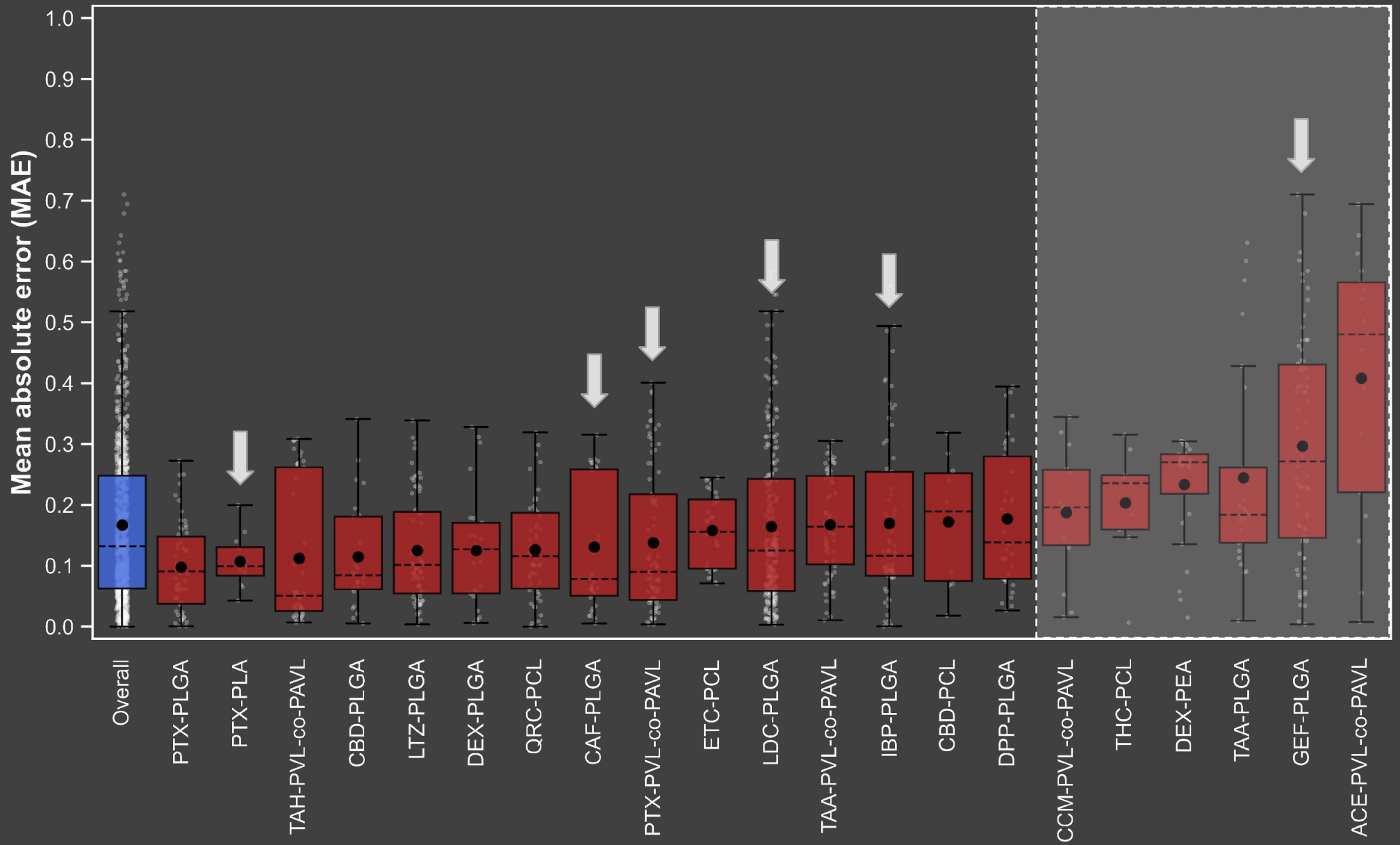


	Random Forest	LightGBM	k-Nearest Neighbors	XGBoost	Decision Tree	Neural Network	Partial Least Squares	Linear Regression
Count	1017	1017	1017	1017	1017	1017	1017	1017
Mean	0.17	0.18	0.17	0.19	0.20	0.23	0.29	0.35
Stdev	0.13	0.14	0.15	0.15	0.16	0.19	0.23	0.27
Median	0.13	0.14	0.14	0.15	0.17	0.18	0.25	0.30

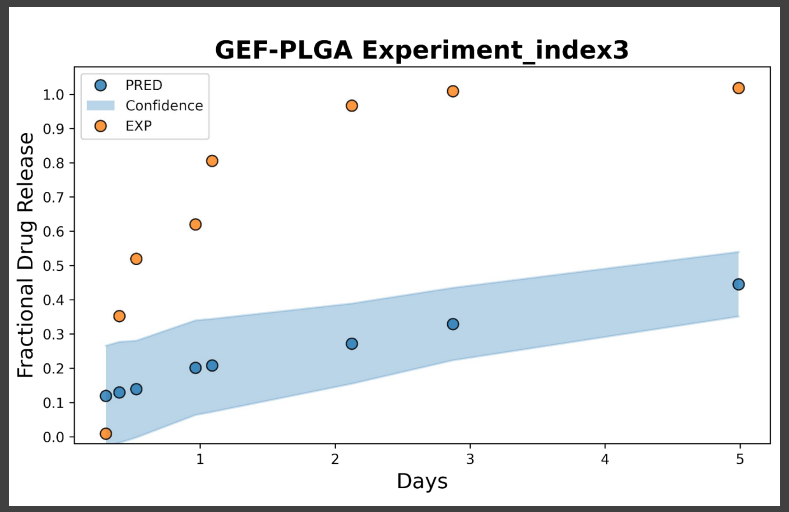
Random Forest Model Performance

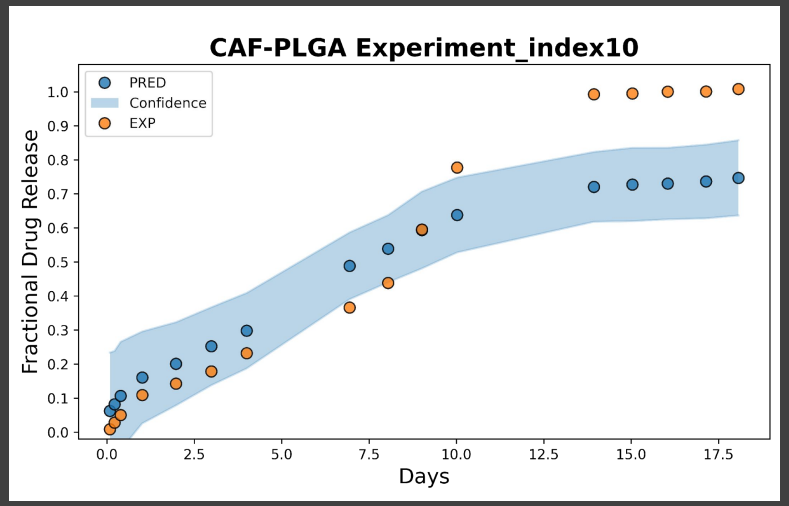


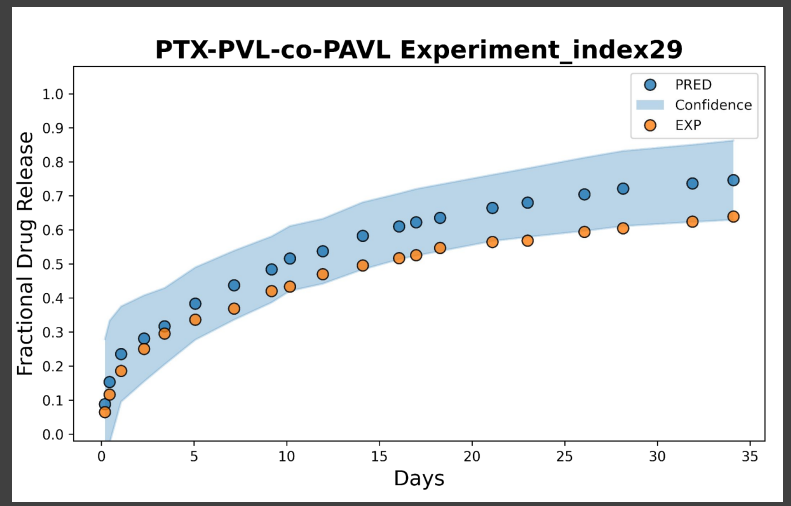
Random Forest Model Performance

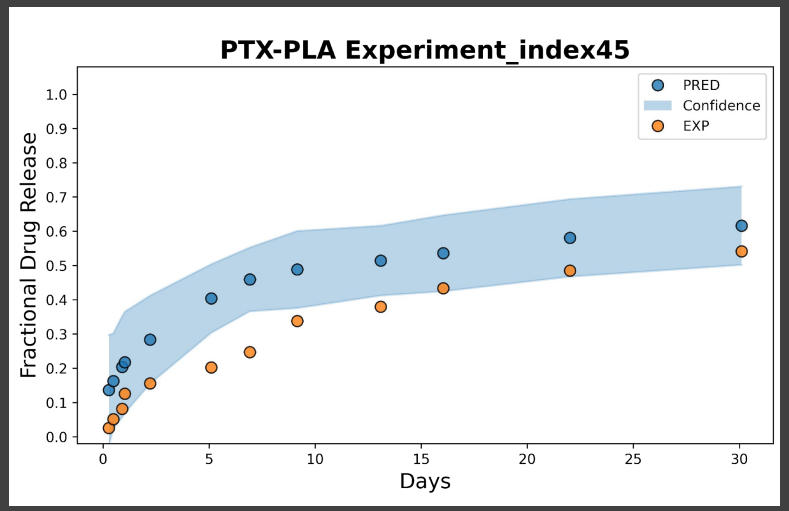


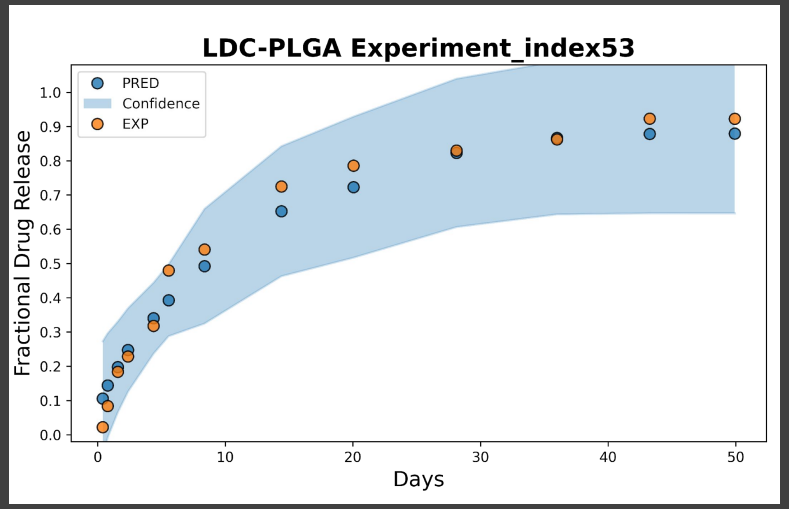
Random Forest Model Predictions

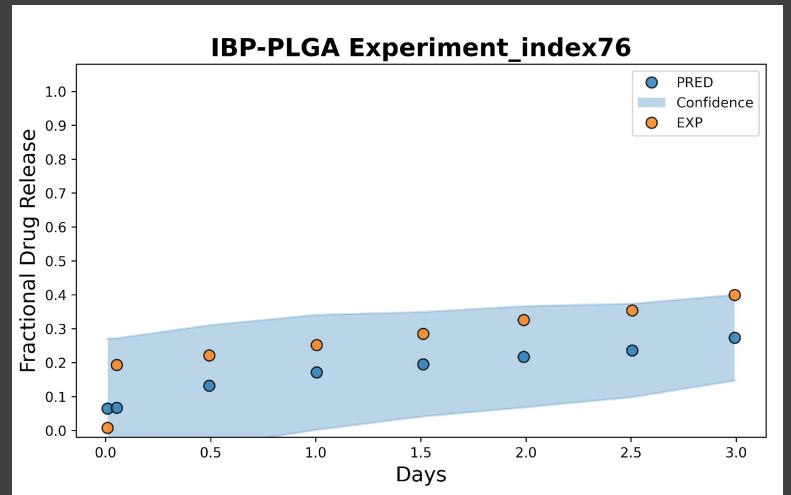








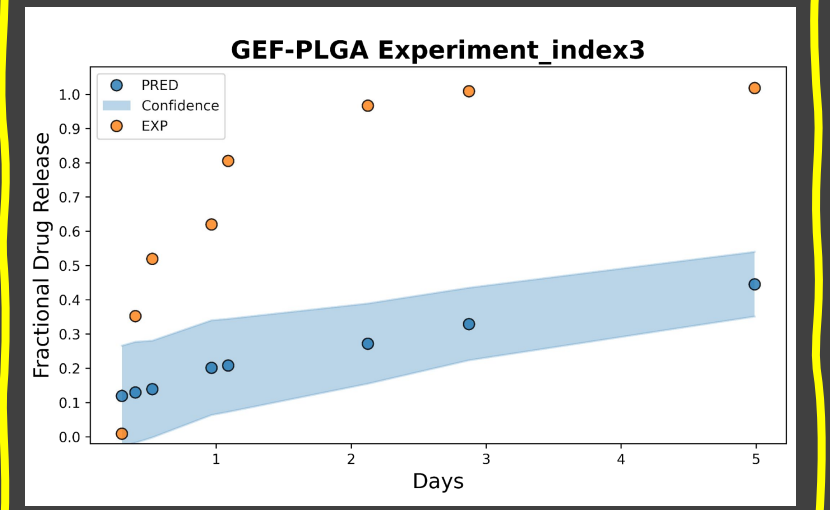


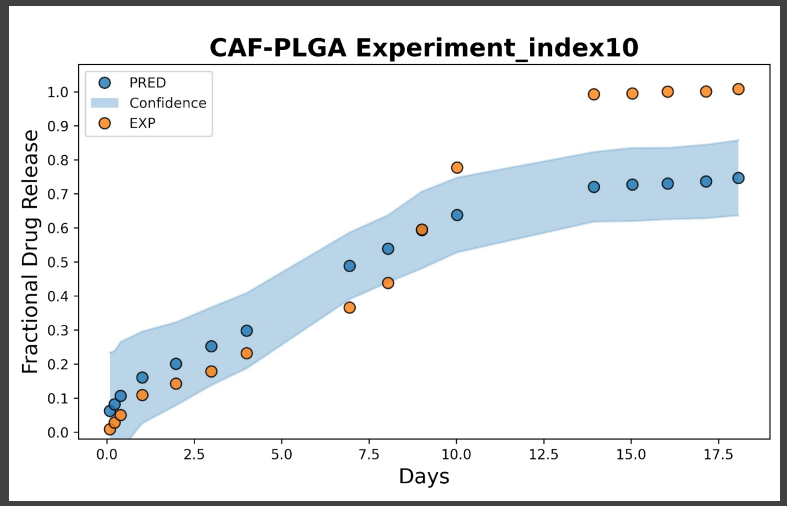


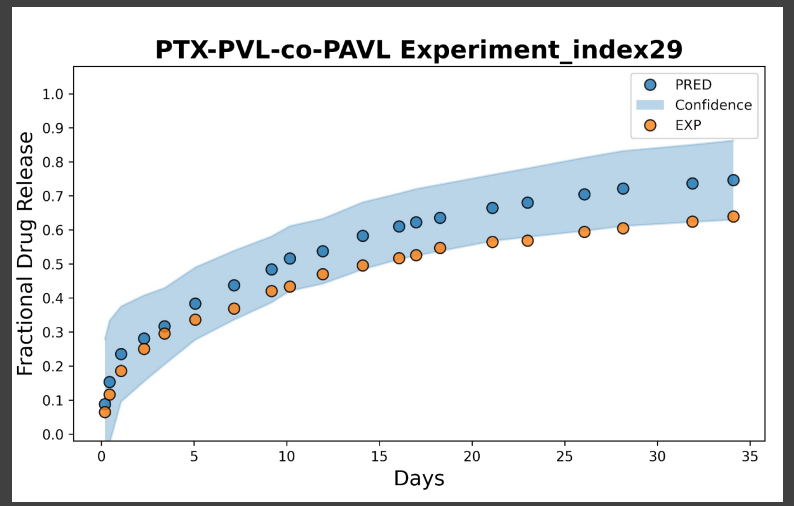
<u>Confidence value calculations:</u>

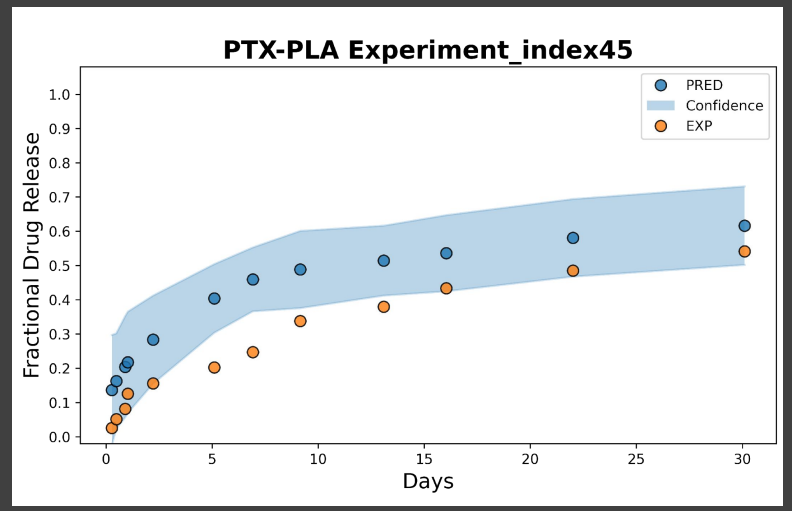
- Polimis et al, (2017), Confidence Intervals for Random Forests in Python, Journal of Open Source Software, 2(19), 124, doi:10.21105/joss.00124

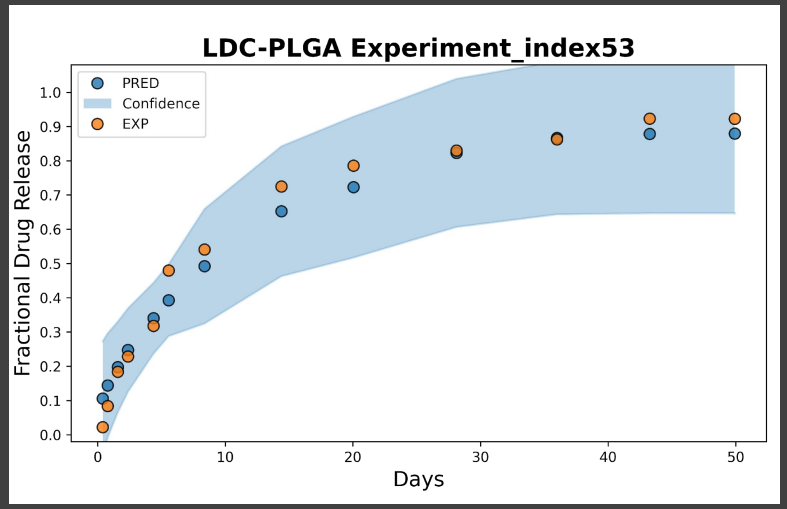
Random Forest Model Predictions

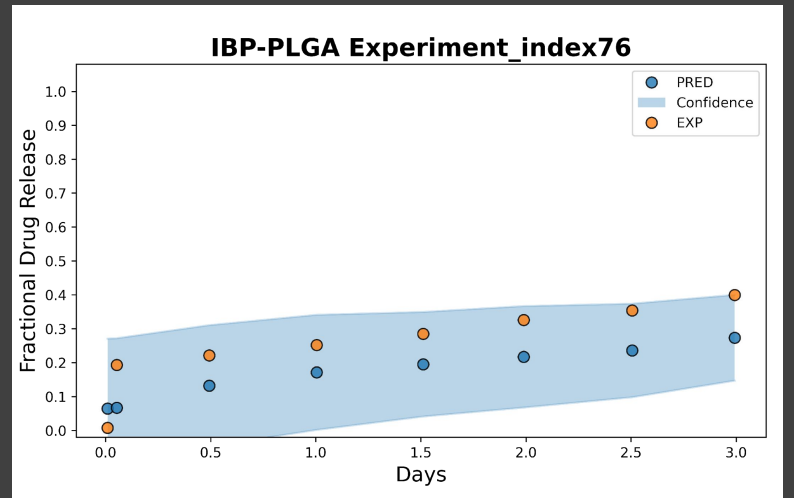








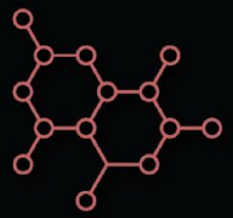




Confidence value calculations:

- Polimis et al, (2017), Confidence Intervals for Random Forests in Python, Journal of Open Source Software, 2(19), 124, doi:10.21105/joss.00124

Analysis of Gefitinib



Gefitinib

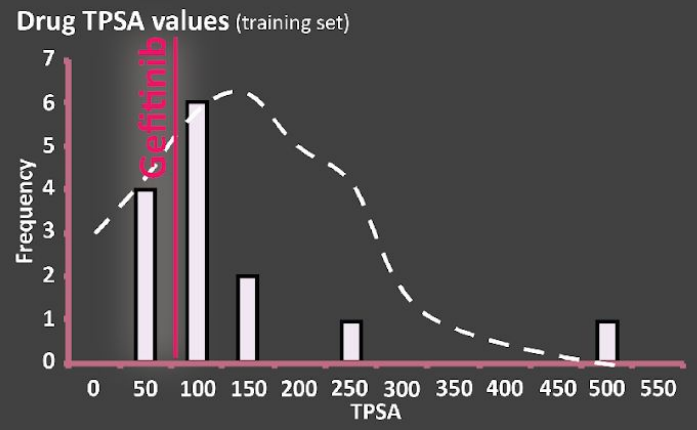
pKa: 5.4

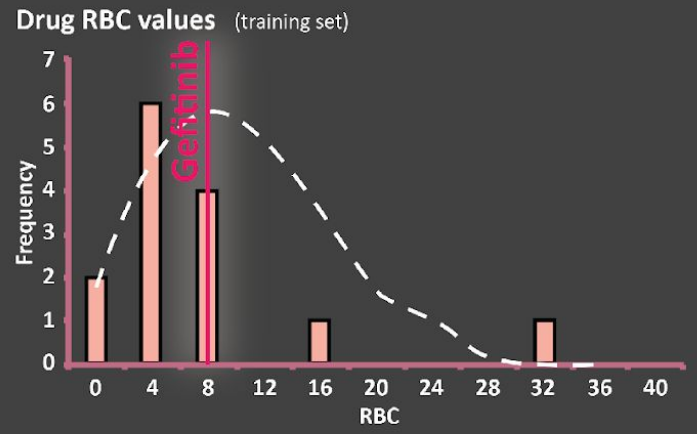
LogP: 3

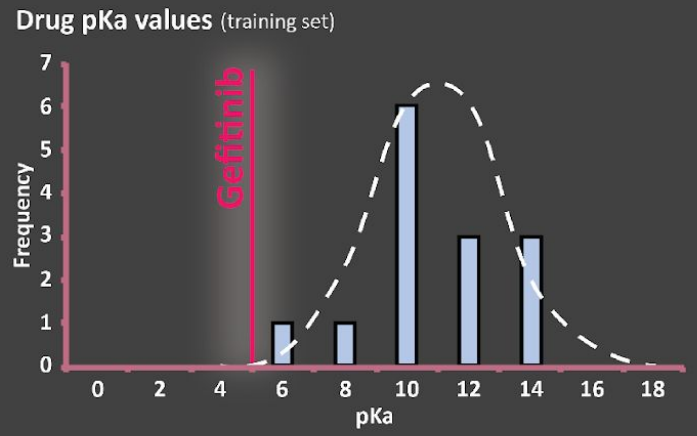
Tm: 213 ℃

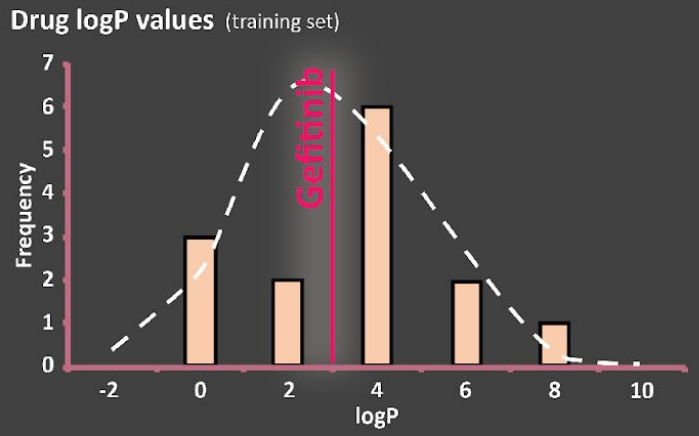
TPSA: 69 Å²

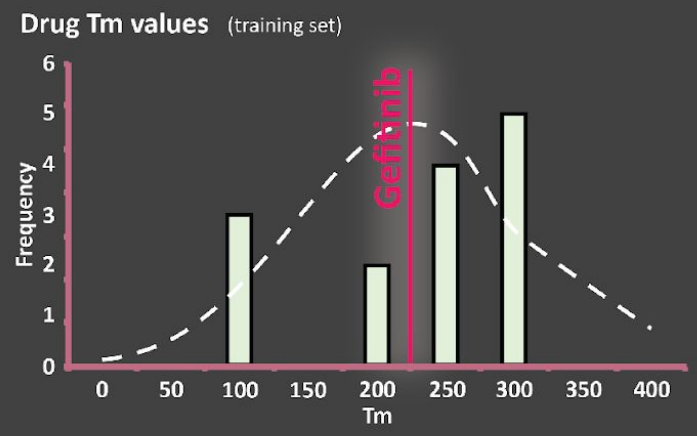
RBC: 8









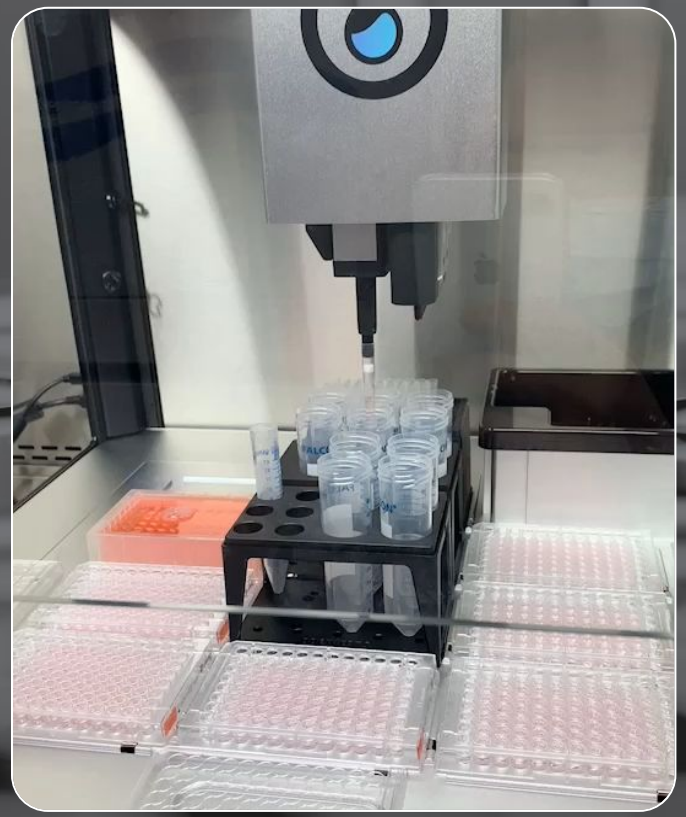


Extracting Data from Previous Studies…

Drug	Polymer Characteristics		Initial Polymer-Drug Ratio (w/w)	Formulation Characteristics		
	Molecular Weight (kDa)	LA:GA		Size (µm)	EE (%)	DL (%)
Lidocaine	N/A	50:50	22:1	14.4	N/A	4
Doxycycline	60	50:50	1.7:1	200 ± 1.52	24.7 ± 0.53	N/A
Etanidazole	40-75	50:50	N/A	422.3 ± 174.5	54.83 ± 0.30	5
Celecoxib	N/A	85:15	N/A	1.11 ± 0.08	51.48 ± 0.42	14.93 ± 0.12
Doxorubicin HCl	10	50:50	7.5:1	3.6 ± 0.4	74.9 ± 3.7	N/A
Simvastatin	50-75	85:15	11.5:1	3.963	96.74 ± 1.46	N/A
Budesonide	75	50:50	10:1	16.67 ± 1.24	74.28 ± 0.83	N/A
Ciprofloxacin	31	50:50	22:1	5.7	8.91 ± 2.90	0.40 ± 0.13
Dexamethasone acetate	19	75:25	N/A	1	N/A	2.5
Leuprolide acetate	N/A	50:50	4:1	N/A	97	20

Experimental Automation

High-throughput HPLC Sample Preparation



In Vitro Drug Combination Studies

Acknowledgments

ALLEN LAB

Christine Allen

Pauric Bannigan

the matter lab

Alán Aspuru-Guzik

Riley Hickman

Matteo Aldeghi

Florian Hase

Pharmacy

Chemistry

Computer Science

UNIVERSITY OF TORONTO

VECTOR INSTITUTE

FOR ARTIFICIAL INTELLIGENCE