

Title	Machine learning methods for prediction of food effects on bioavailability: A comparison of Support Vector Machines and Artificial Neural Networks
Authors	Bennett-Lenane, Harriet;Griffin, Brendan T.;O'Shea, Joseph P.
Publication date	2021-09-24
Original Citation	Bennett-Lenane, H., Griffin, B. T. and O'Shea, J. P. (2021) 'Machine learning methods for prediction of food effects on bioavailability: A comparison of support vector machines and artificial neural networks', European Journal of Pharmaceutical Sciences. https://doi.org/10.1016/j.ejps.2021.106018
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1016/j.ejps.2021.106018
Rights	© 2021, the Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/) - https://creativecommons.org/licenses/by/4.0/
Download date	2026-08-31 13:15:08
Item downloaded from	https://hdl.handle.net/10468/12059



UCC

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Supplementary Materials

Machine Learning Methods for Prediction of Food Effects on Bioavailability: A Comparison of Support Vector Machines and Artificial Neural Networks

Harriet Bennett-Lenane, Brendan T. Griffin, Joseph P. O'Shea.

School of Pharmacy, University College Cork, Cork, Ireland.

*Correspondence to: Joseph O'Shea; Tel.: +353 (0)21 4901665, Email address: joseph.oshea@ucc.ie

1. Results of Statistical Analysis comparing Positive, Negative and No FE groups for drugs licensed from 2016-2020
2. Visual comparison of the sensitivity, precision, specificity and Matthews correlation coefficient (MCC) performance metrics calculated for the test set of the optimum support vector machine (SVM) and artificial neural network (ANN) models produced in this study to predict FE classification.

1. Results of Statistical Analysis comparing Positive, Negative and No Food Effect (No FE) groups for drugs licensed from 2016-2020

Drug Property	Descriptors	Negative	No FE	Positive	Statistical Tests	Neg vs No FE	Neg vs Pos	No FE vs Pos
S+logP	<i>n</i>	17	80	44	Levene's Test	0.08	0.15	0.81
	<i>Median</i>	2.09	2.80	3.61	Bootstrap/t-test	0.018 ^B	0.002 ^B	0.021 ^t
	<i>Mean</i>	1.50	2.90	3.67	Mean Difference	-1.41	-2.17	-0.76
	<i>SD of Mean</i>	2.29	1.72	1.77	95% Confidence Interval	(L) -2.56 (U) -0.34	(L) -3.41 (U) -0.98	(L) 1.41 (U) -0.11
	<i>Q1, Q3,</i>	-0.45, 3.33	2.02, 3.90	2.71, 4.80				
	<i>Min, Max</i>	-2.43, 4.21	-1.38, 7.49	-1.34, 7.25				
	<i>Variance</i>	5.26	2.97	3.13				
Hydrogen Bond Donors	<i>n</i>	17	80	44	Levene's Test	0.075	0.26	0.34
	<i>Median</i>	4	2	2	Bootstrap	0.00	0.00	0.86
	<i>Mean</i>	3.53	1.90	1.86	Mean Difference	1.63	1.67	0.04
	<i>SD of Mean</i>	1.55	1.09	1.15	95% Confidence Interval	(L) 0.85 (U) 2.39	(L) 0.86 (U) 2.48	(L) -0.36 (U) 0.46
	<i>Q1, Q3</i>	2.5, 4.5	1, 3	1, 3				
	<i>Min, Max</i>	0, 6	0, 4	0, 4				
	<i>Variance</i>	2.39	1.18	1.33				
Hydrogen Bond Acceptors	<i>n</i>	17	80	44	Levene's Test	0.60	0.42	0.04
	<i>Median</i>	8	6	7	Bootstrap	0.06	0.67	0.09
	<i>Mean</i>	7.71	6.28	7.30	Mean Difference	1.43	0.41	-1.02
	<i>SD of Mean</i>	2.69	2.57	3.49	95% Confidence Interval	(L) -0.04 (U) 2.91	(L) -1.47 (U) 2.30	(L) -2.22 (U) 0.18
	<i>Q1, Q3</i>	5.5, 10	5, 7.75	5, 9				
	<i>Min, Max</i>	4, 13	1, 15	2, 16				
	<i>Variance</i>	7.221	6.61	12.17				
logD_{7.4}	<i>n</i>	17	80	44	Levene's Test	0.05	0.16	0.73
	<i>Median</i>	1.12	2.18	2.93	Bootstrap/t-test	0.09 ^B	0.003 ^B	0.003 ^t
	<i>Mean</i>	0.82	1.98	3.01	Mean Difference	-1.16	-2.19	-1.03
	<i>SD of Mean</i>	2.65	1.78	1.93	95% Confidence Interval	(L) -2.65 (U) 0.22	(L) -3.66 (U) -0.80	(L) -1.71 (U) -0.35
	<i>Q1, Q3</i>	-0.57, 2.93	0.99, 3.13	1.75, 4.17				
	<i>Min, Max</i>	-5.18, 4.21	-2.87, 6.64	-2.26, 7.25				
	<i>Variance</i>	7.07	3.17	3.71				
Polar Surface Area	<i>n</i>	17	80	44	Levene's Test	0.16	0.87	0.05
	<i>Median</i>	107.41	90.37	100.01	Bootstrap	0.00	0.07	0.047
	<i>Mean</i>	125.09	89.12	104.26	Mean Difference	35.92	20.77	-15.15
	<i>SD of Mean</i>	35.87	33.26	43.93	95% Confidence Interval	(L) 18.03 (U) 54.07	(L) -1.06 (U) 42.94	(L) -30.40 (U) 0.19
	<i>Q1, Q3</i>	92.88, 157.65	70.47, 108.48	77.36, 132.71				
	<i>Min, Max</i>	76.10, 178.36	22.00, 188.80	29.46, 199.58				
	<i>Variance</i>	1286.99	1105.95	1930.20				
Dose (mg)	<i>n</i>	17	80	44	Levene's Test	0.25	0.009	0.258
	<i>Median</i>	80	50.00	200.00	Bootstrap	0.64	0.001	0.005
	<i>Mean</i>	105.36	121.15	242.57	Mean Difference	-15.79	-137.20	-121.41
	<i>SD of Mean</i>	96.00	205.71	193.88	95% Confidence Interval	(L) -79.20 (U) 48.38	(L) -212.56 (U) -62.23	(L) -195.66 (U) -44.74
	<i>Q1, Q3</i>	25.00, 175.00	10.00, 143.75	100.00, 375.00				

Drug Property	Descriptors	Negative	No FE	Positive	Statistical Tests	Neg vs No FE	Neg vs Pos	No FE vs Pos
	<i>Min, Max</i>	1.00, 300.00	0.20, 1340.00	2.00, 800.00				
	<i>Variance</i>	9216.31	42314.80	37588.93				
S+Sw	<i>n</i>	17	80	44	Levene's Test	0.18	0.18	0.86
	<i>Median</i>	-0.49	-0.99	-1.54	Bootstrap/t-test	0.13B	0.003B	0.002t
	<i>Mean</i>	-0.38	-0.92	-1.56	Mean Difference	0.54	1.18	0.64
	<i>SD</i>	1.39	1.07	1.05	95% Confidence Interval	(L) -0.10	(L) 0.51	(L) 0.24
	<i>Q1, Q3</i>	-1.49, 0.22	-1.55, -0.31	-2.35, -0.76		(U) 1.25	(U) 1.91	(U) 1.03
	<i>Min, Max</i>	-2.05, 2.98	-3.61, 2.02	-4.42, 1.01				
	<i>Variance</i>	1.94	1.14	1.097				
MAD	<i>n</i>	17	80	44	Levene's Test	0.07	0.07	0.78
	<i>Median</i>	2.46	2.78	1.02	t-test	0.96	0.06	0.001
	<i>Mean</i>	2.77	2.75	2.11	Mean Difference	0.15	0.67	0.64
	<i>SD</i>	1.54	1.06	1.02	95% Confidence Interval	(L) -0.60	(L) -0.02	(L) 0.25
	<i>Q1, Q3</i>	1.57, 3.54	2.15, 3.38	1.26, 2.71		(U) 0.63	(U) 1.33	(U) 1.03
	<i>Min, Max</i>	0.57, 6.14	0.02, 5.84	-0.13, 4.34				
	<i>Variance</i>	2.36	1.13	1.04				
Molecular Weight (g/mol)	<i>n</i>	17	80	44	Levene's Test	0.09	0.86	0.04
	<i>Median</i>	457.69	426.31	446.95	Bootstrap	0.47	0.41	0.02
	<i>Mean</i>	453.74	419.89	497.65	Mean Difference	33.85	-43.91	-77.78
	<i>SD of Mean</i>	184.44	134.86	180.56	95% Confidence Interval	(L) -59.07	(L) -147.60	(L) -141.53
	<i>Q1, Q3</i>	302.62, 570.61	348.99, 491.83	368.95, 557.87		(U) 126.99	(U) 60.81	(U) -22.05
	<i>Min, Max</i>	146.24, 804.04	109.13, 889.02	287.22, 1113.21				
	<i>Variance</i>	34018.56	18188.43	32601.00				
Dose/Solubility Ratio	<i>n</i>	17	80	44	Levene's Test	0.57	0.26	0.28
	<i>Median</i>	2.14	2.47	3.86	t-test	0.28	0.00	0.00
	<i>Mean</i>	2.10	2.47	3.71	Mean Difference	-0.37	-1.61	-1.24
	<i>SD of Mean</i>	1.36	1.23	1.20	95% Confidence Interval	(L) -1.03	(L) -2.32	(L) -1.70
	<i>Q1, Q3</i>	1.08, 3.22	1.62, 3.26	3.32, 4.40		(U) 0.30	(U) -0.90	(U) -0.79
	<i>Min, Max</i>	-0.89, 4.05	-0.54, 5.83	0.18, 7.02				
	<i>Variance</i>	1.86	1.52	1.43				
Rotatable Bonds	<i>n</i>	17	80	44	Levene's Test	0.15	0.89	0.02
	<i>Median</i>	6	5	5	Bootstrap	0.26	0.96	0.20
	<i>Mean</i>	6.65	5.50	6.60	Mean Difference	1.15	0.079	-1.07
	<i>SD of Mean</i>	5.26	3.41	4.97	95% Confidence Interval	(L) -0.86	(L) -2.81	(L) -2.79
	<i>Q1, Q3</i>	2.50, 8.50	3.23, 7.75	3.00, 9.00		(U) 3.15	(U) 2.97	(U) 0.49
	<i>Min, Max</i>	0.00, 22.00	0.00, 15.00	0.00, 24.00				
	<i>Variance</i>	27.62	11.60	24.72				

2. Visual comparison of the sensitivity, precision, specificity and **Matthews Correlation Coefficient** (MCC) performance metrics calculated for the test set of the optimum **Support Vector Machines** (SVM) and **Artificial Neural Network** (ANN) models produced in this study to predict **food effect** (FE) classification.

