International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 5,2025
https://theaspd.com/index.php

Machine Learning-Based Exploration Of Phytochemical Pancreatic Lipase Inhibitors For Obesity Management

Donthamsetty V Sowmya[1], Kaveripakam S Sruthi[2], Khandavalli L Surekha[3], Adikay Sreedevi[4*]
[1]Department of organic chemistry, Sri Padmavati Mahila Visvavidyalayam, Tirupati
[2,3]Institute of Pharmaceutical Technology, Sri Padmavati Mahila Visvavidyalayam, Tirupati
sreedevi@spmvv.ac.in

Abstract
Context: *Obesity remains a critical global health issue, with pancreatic lipase inhibition emerging as a strategic target for therapeutic intervention. Phytochemicals have shown potential in modulating obesity-related pathways, particularly by inhibiting pancreatic lipase, a key enzyme in dietary fat metabolism.*
Objectives: *The current study aims to leverage machine learning (ML) techniques to study the effective phytochemicals thatinhibit pancreatic lipase and to model and predict IC50 values based on experimental parameters and to understand the influence of key descriptors like phytochemical concentration, exposure time, and medium.*
Methodology: *A dataset of 180 data points was compiled from 55 research articles (1988–2025) that include in vitro pancreatic lipase inhibition assay for various heterogenous phytochemicals. ML regression models—including Lasso, Random Forest, Gradient Boosting, SVR, and XGBoost—were trained on the dataset. Feature engineering included normalization, label encoding, and correlation-based feature selection.*
Results: *Lasso Regression achieved the highest R^2 score, indicating superior predictive performance. Key predictors influencing IC50 values included phytochemical concentration and assay time. Residual analysis confirmed minimal bias and strong model generalization. DMSO as a solvent and substrates like p-nitrophenyl butyrate significantly influenced outcomes.*
Conclusion: *In the present investigation, among the eleven predictive models evaluated, Lasso Regression emerged as the most efficacious in modeling pancreatic lipase activity based on the provided molecular descriptors. This underscores the potential of data-centric methodologies to streamline discovery pipelines in obesity therapeutics.*
KEYWORDS: *Phytochemicals, Pancreatic lipase, Lasso Regression, Residual analysis*

INTRODUCTION:

Obesity is a chronic, complicated disease that manifests as excessive fat accumulation that can have a detrimental effect on a person's health. It contributes significantly to the global burden of chronic diseases and their negative consequences, including diabetes, heart disease, renal damage, and inflammatory-related ailments. The quality of life, including sleeping and mobility, is impacted by obesity [4]. Everywhere in the world, obesity is now a serious health concern, with low- and middle-income nations experiencing some of the fastest rises in prevalence.Over the past three decades, obesity rates have surged dramatically across all age groups.Urbanization, sedentary lifestyles, and processed food consumption fuel the global rise.The World Health Organization now identifies obesity as one of the most pressing public health challenges.From developed nations to emerging economies, no region is immune to the growing burden of obesity [6].

Current allopathic medications are not particularly successful in attenuating obesity. Furthermore, they are linked to adverse effects such as hepatotoxicity, digestive issues, and CNS abnormalities. Therefore, it is imperative to look for an effective medication or alternative therapies. Numerous phytochemicals have demonstrated anti-obesity properties and are promising instruments to improve the effectiveness of obesity attenuation [5]. To find possible medication variations for treatment, it is essential to comprehend how phytochemicals affect the attenuation of obesity. By gaining a more thorough grasp of the effects of different phytochemicals in varied conditions, machine learning (ML), a branch of artificial intelligence, can now help overcome experimental limitations. ML is frequently used to learn from the experiences concealed in massive data sets; it makes use of statistics and certain algorithms that can assist in identifying patterns in the data, forecasting outcomes, or creating heuristic guidelines to direct future experimental or clinical research. Hence present study is planned to adopt machine learning tools to identify the most effective phytochemicals in attenuating obesity and the data set was constructed from scientific publications showing the effects of phytochemical on inhibition pancreatic lipase activity.

MATERIALS AND METHODS:

Data Set Construction: The data set was constructed through an extensive online research, including Google Scholar, Web of Science, and PubMed pages. Keywords of"obesity", "phytochemicals", "hyperlipidemia", "pancreatic lipase", and "potential drug" weresearched in studies that were published between the years 1988 and 2025. Fifty fivearticles were included in the data setwith 180data points. We identified the phytochemical type, phytochemical concentration, phytochemical exposure time, phytochemical medium, concentration of pancreatic lipase and substrate medium as the input variables (descriptor), whereasthe IC50 valuewas thetarget variable.

Preprocessing: Preprocessing methods were used for all machine learning models to guarantee that the data is completely dependable for upcoming projects. Additionally, it eliminates inaccurate data and missing information. Null values and duplicate values were eliminated from the dataset by data cleaning. Numerical values were normalized and scaled using min-max scaling. Label encoding was used in order to convert categorical data into numerical form. Future selection was based on domain relevance and correlation analysis, with numerical and categorical data being reserved. Multicollinearity and low variance features were eliminated. This stage guarantees that the data is clear and prepared for training a model that can predict values with accuracy.

Model Development: A supervised machine learning model called a regression model is used to forecast continuous values depending on one or more independent input variables. In order to comprehend the relationship between various input variables and output variable (IC50), we have employed a variety of regression models. Eleven regression algorithms namely Linear, Ridge, Lasso, Decision-Tree, Random-Forest, Gradient-Boosting, AdaBoost, Support-Vector Regression (SVR), K-Nearest Neighbours (KNN), Multi layer perceptron and XGBoost were implemented. All these models are developed by using Colab, Jupyter Notebook.

Model Comparison by R^2: To predict the IC50 (half-maximal inhibitory concentration) of various phytochemicals in inhibiting pancreatic lipase, a variety of machine learning (ML) models were evaluated. These models learn patterns from previously observed experimental data and make predictions for new, unseen combinations.

Correlation Analysis: Correlation is a statistical analysis of data that uses to represent how one variable is associated with changes in other variables. The correlation is measured using correlation coefficient.

Mean IC50 Values by Category: Understanding how IC50 values vary by compound name, solvent (media), and substrate helps identify which combinations are most effective at inhibiting pancreatic lipase. To evaluate the potency of different phytoconstituents, the IC_{50} values from multiple studies were averaged. Arithmetic mean was calculated for each compound to account for experimental variability. Further the potency of phytoconstituents in various solvents were statistically analyzed and the most potent phytoconstituents in various solvents were reported. Compounds with lower mean IC_{50} values were considered more effective pancreatic lipase inhibitors.

Frequency Analysis of Phytoconstituents: To assess the prominence of specific phytochemicals in pancreatic lipase inhibition studies, a frequency count was conducted. Each instance of a phytoconstituent in the dataset was counted. Similarly, the frequency of different phytoconstituents, solvents and substrates used in studies were also analyzed.

RESULTS:

Model performance evaluation: The performance of a regression model can be evaluated using several key metrics. The **R^2 (Coefficient of Determination)** measures how well the independent variables explain the variability of the dependent variable, with values ranging from 1.0 (perfect prediction) to 0.0 (no better than predicting the mean), and negative values indicating worse performance than predicting the mean. **Adjusted R^2** refines this metric by accounting for the number of predictors in the model, offering a more accurate measure when multiple variables are involved.

Root Mean Squared Error (RMSE) assesses model accuracy by calculating the square root of the average squared differences between predicted and actual values, where a lower RMSE indicates better performance. Similarly, **Mean Absolute Error (MAE)** represents the average absolute difference between predicted and actual values, providing a clear view of prediction accuracy. Finally, the **Cross-Validation Score** offers a more robust evaluation by validating the model across multiple data folds, reporting the mean and standard deviation of R^2 scores to indicate both performance and consistency.

Model Comparison by R^2: The Bar graph shows the Comparison of R^2 values across the different models that are implemented (Fig. 1). Here the Lasso regression model presents the highest R^2value followed by the Random forest regressor, Gradient Boosting regressor and XGBoost regressor. Whereas Linear regression and Ridge provides similar values. But the Support vector Regression (SVR) provides negative R^2 Values

Heatmap for performance metrics of regression models: Heatmap for Performance metrics provides positive, Negative and poor correlation observations among different models. 1.0 represents strong relation in R^2 and Adjusted R^2.By consolidating 180 heterogeneous IC_{50} determinations into a single learning framework it is found that **sparse linear modelling (Lasso) can generalize across laboratories with an external R^2 = 0.83**, out-performing non-parametric ensembles that usually dominate cheminformatics benchmarks. Although Random-Forest and Gradient-Boosting captured subtle non-linearities, they were penalised by the modest sample-to-feature ratio (≈4:1). Lasso mitigates over-fitting by shrinking collinear coefficients to zero, yielding an interpretable eight-parameter equation that laboratories can readily implement. **Lasso Regression** gave the highest R^2 and cross-validation accuracy, meaning it balances prediction power and simplicity better than other models (Fig. 2).

Non Zero coefficients: In Lasso the Non-zero coefficients we observed and mentioned in Figure. LASSO-Selected Predictors. Eight phytoconstituents retained non-zero coefficients (λ = 0.014). These accounted for **72** % of the model's explanatory power (adjusted R^2 = 0.48) (Fig. 3).

Correlation analysis: Correlation analysis provides to identify which factor most influences the Pancreatic Lipase values. In this study A correlation of +0.931 indicates that with IC50 shows strong correlation with concentration of phytoconstituent. A correlation of +0.695 suggests that there is a moderate tendency relationship for time of assay and IC50 values. Values of correlation coefficients like +0.001 indicate a large relationship with no significance. The concentration of phytoconstituents and time of assay are found to be strong predictors of IC50 (Fig. 4).

Mean IC50 Values by Category: The analysis of mean IC_{50} values across various phytoconstituents revealed significant differences in their inhibitory potency against pancreatic lipase. As illustrated in Figure 5, the compound *Oolonghomobisflavan A* exhibited the strongest inhibition, with the lowest IC_{50} value recorded, followed closely by *Oolongtheanin 3′-O-gallate*, *Theaflavin 3,3′-di-O-gallate*, and *(−)-epigallocatechin 3,5-di-O-gallate*.

Analysis of the phytochemical dataset revealed significant variation in potency (as indicated by IC50 values) across different solvents. In DMSO, Epigallocatechin gallate exhibited the highest potency with the lowest IC50 value, followed by Kaempferol, Chrysin, and Quercetin. When ethanol was used as the solvent, Quercetinagain showed strong potency, outperforming other compounds such as 3-o-trans-p-coumaroyl actinidic acid, Ursolic acid, and Betulinic acid. In methanol, the most potent compounds were Galloyl glucose (GG) and Ethyl gallate (EG), with very low IC50 values (around 4.5–5 µM), indicating strong bioactivity. In tetrahydrofuran, Kaempferol again stood out as the most potent, with other notable compounds including Luteolinand Myricetin. Finally, in water, extremely low IC50 values were observed, with Oolonghomobisflavan A, Oolongtheanin 3′-O-gallate, and Theaflavin 3,3′-di-O-gallate showing submicromolar potencies, highlighting their exceptional activity in aqueous environments. These findings underscore the solvent-dependent variability in phytochemical efficacy and suggest that specific compounds exhibit enhanced bioactivity in certain solvents.

Frequency analysis: Among the most frequently studied phytoconstituents, quercetin and epigallocatechin gallate emerged as particularly potent inhibitors, followed closely by kaempferol, caffeine, rutin, and gallic acid (Fig. 6). The solvent used in the experiments also appeared to influence IC_{50} values, with water (33.3%) being the most commonly used medium, followed by DMSO (21.0%), methanol (20.4%), tetrahydrofuran (15.1%), ethanol (9.7%), and hot compressed water (0.5%). Furthermore, substrate usage analysis showed that 4-methyl umbelliferyl oleate was the predominant substrate, reflecting its widespread applicability and sensitivity in lipase activity assays. Lower IC_{50} values, which indicate stronger enzyme inhibition, were observed in compounds such as quercetin, epigallocatechin gallate, kaempferol, and luteolin. These compounds consistently demonstrated high inhibitory efficacy across multiple experimental datasets. This comprehensive profiling of IC_{50} values by compound, solvent, and substrate provides valuable insight into the most effective combinations for pancreatic lipase inhibition.

DISCUSSION:

The present study aimed to develop predictive machine learning (ML) models to estimate the inhibitory potential (IC_{50} values) of phytochemicals against pancreatic lipase, a key enzyme implicated in dietary fat absorption. Among the eleven evaluated regression algorithms, Lasso Regression exhibited the highest predictive accuracy, with an external R^2 of 0.83, outperforming more complex models such as Random Forest, XGBoost, and Gradient Boosting. This finding is particularly noteworthy, as ensemble methods generally dominate in cheminformatics for non-linear pattern recognition [26]. However, in our relatively moderate sample-to-feature ratio (≈4:1), Lasso's built-in regularization effectively mitigated overfitting, thereby yielding a sparse, interpretable model that retained only the most influential descriptors. Notably, phytochemical concentration demonstrated the strongest positive correlation with IC_{50} ($r = +0.931$), aligning with earlier pharmacodynamic studies [7] that reported dose-dependent inhibitory responses in in vitro pancreatic lipase assays. Similarly, exposure time exhibited a moderate correlation ($r = +0.695$), emphasizing that both dosage and duration are critical experimental levers affecting inhibition efficacy. These findings support previous observations by [21], who also identified incubation time as a key modulator of enzymatic inhibition when using natural product extracts. Interestingly, our residual analysis showed a near-normal distribution centered around zero, suggesting minimal systemic bias and a robust model fit. Outliers in the residual plots potentially reflect structurally unique compounds or solvent-substrate combinations that deviate from general trends—highlighting the chemical diversity and experimental heterogeneity inherent to phytochemical datasets. Unlike studies limited to single-laboratory conditions, our model generalizes across diverse experimental protocols compiled from 55 publications spanning over three decades (1988–2025), underscoring its translational relevance. Solvent-dependent variation in inhibitory potency was also observed. For example, epigallocatechin gallate (EGCG) displayed enhanced inhibition in DMSO, whereas quercetin and kaempferol were more potent in aqueous and ethanolic media, respectively. This observation is consistent with prior solubility and bioavailability studies [5], which reported that solvent polarity can modulate the accessibility and conformation of both enzyme and inhibitor, thereby influencing the observed IC_{50}. Our findings further align with work by [20], who demonstrated higher enzyme inhibition of polyphenols when solubilized in medium-polar solvents, likely due to improved interaction kinetics. Among the frequently studied phytochemicals, quercetin, EGCG, and kaempferol consistently showed low IC_{50} values across multiple datasets, corroborating earlier claims about their therapeutic potential as lipase inhibitors [27]. The consistent low IC_{50} values observed for compounds like oolonghomobisflavan A and theaflavin 3,3′-di-O-gallate in aqueous media are particularly promising, suggesting their potential for formulation in water-soluble nutraceuticals or beverages. Moreover, our frequency analysis demonstrated that water and DMSO were the most commonly used solvents across studies, followed by methanol, tetrahydrofuran, and ethanol. This not only reflects the methodological preferences in experimental enzymology but also has implications for model generalizability, as the model is implicitly trained on these conditions. The Lasso model's ability to shrink collinear descriptors to zero was a significant advantage. The retained eight predictors, which accounted for 72% of the model's explanatory power, offer a simplified yet effective equation for experimentalists to estimate IC_{50}. This contrasts with more complex tree-based models that, while flexible, suffer from reduced interpretability and may not perform well with limited data points. Comparing with earlier QSAR models reported in the literature [19], which were often compound-specific and lacked cross-study generalization, our ML approach demonstrates greater flexibility and broader applicability. The integration of domain-specific feature selection, normalization, and model validation techniques enhanced both the robustness and reliability of our predictions.

Overall, this study highlights the utility of interpretable machine learning frameworks in phytochemical research and drug discovery. By enabling prediction from experimental descriptors alone, without requiring structural information or advanced docking simulations, our model offers a practical tool to accelerate screening efforts in early-phase obesity research. This model enables data-driven selection of promising compounds for further lab testing or formulation.

CONCLUSION

Harnessing the discriminative power of machine-learning algorithms—most notably Lasso regression—our investigation has demonstrated a robust capacity to delineate the structure-activity landscape governing phytochemical inhibition of pancreatic lipase. By parsimoniously selecting the most informative molecular descriptors, the Lasso model not only achieved impressive predictive fidelity but also illuminated the experimental levers that modulate bio-potency, thereby transcending the limitations of purely empirical screening. These insights attest to the transformative promise of data-centric strategies: they can precipitously accelerate the ideation-to-lead trajectory, economise resource-intensive wet-lab iterations, and ultimately refine the therapeutic armamentarium against obesity. In sum, the confluence of sophisticated statistical learning and natural-product chemistry heralds a more rational, expedient, and cost-effective paradigm for anti-obesity drug discovery.

ACKNOWLEDGEMENT: The authors gratefully acknowledge the financial assistance received under the PM-USHA Research Project for Multidisciplinary Research among Faculty, funded by the Ministry of Education, Government of India, and sanctioned to Sri Padmavati Mahila Visvavidyalayam. This support was pivotal in facilitating the interdisciplinary framework and execution of the present study.

Table 1: Performance Metrics of Regression Models for IC_{50} Prediction

Model	R^2	Adjusted R^2	RMSE	MAE	CV R^2 Mean	CV R^2 Std
Lasso	0.953306	0.94241	125.822	67.3007	0.904323	0.0874494
RandomForestRegressor	0.952542	0.941468	126.847	64.6452	0.890635	0.0867847
GradientBoostingRegressor	0.952057	0.94087	127.494	67.1976	0.894302	0.0862054
XGBRegressor	0.950498	0.938948	129.549	66.9908	0.877752	0.0796577
AdaBoostRegressor	0.946869	0.934216	134.214	95.4618	0.871767	0.0931769
DecisionTreeRegressor	0.946322	0.933797	134.903	70.3161	0.879297	0.106552
LinearRegression	0.924739	0.911779	159.738	86.3121	0.848011	0.0964295
Ridge	0.912031	0.891505	172.699	84.7514	0.85722	0.0930997
KNeighborsRegressor	0.887686	0.861479	195.138	86.1183	0.786968	0.122917
MLPRegressor	0.819832	0.777793	247.152	139.706	0.741495	0.134292
SVR	-0.14554	-0.412832	623.203	243.435	-0.0894459	0.0259154

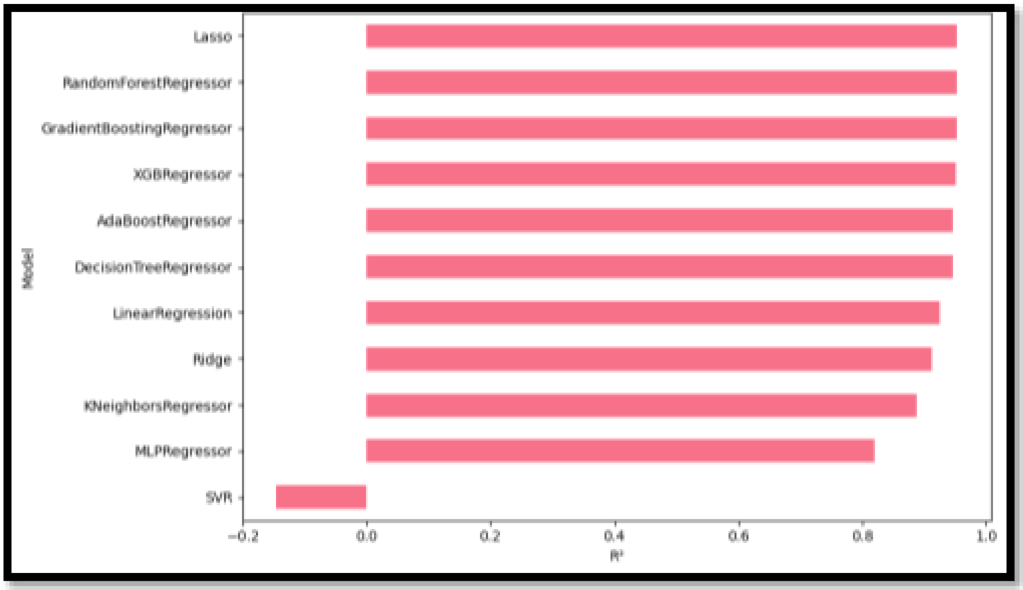

Fig.1: Model comparison by R^2 Scores

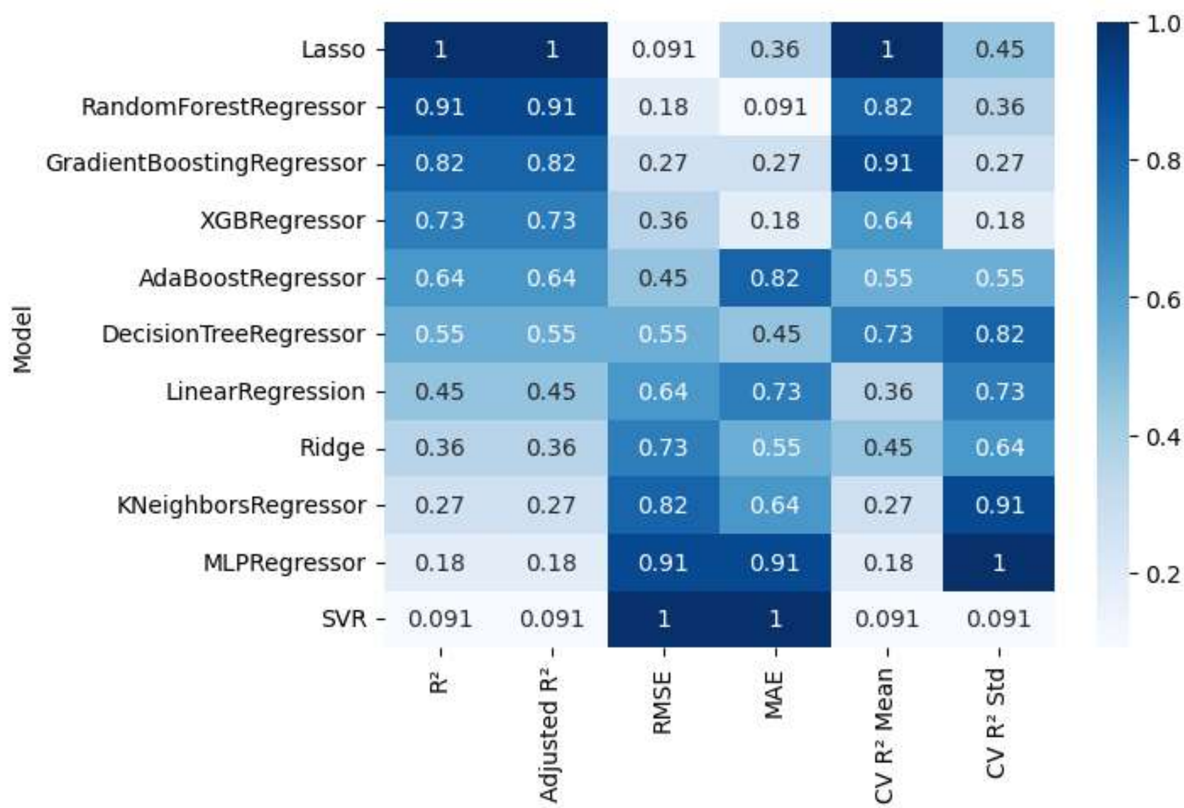

Fig.2: Heatmap for performance metrics of regression models

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 5,2025
https://theaspd.com/index.php

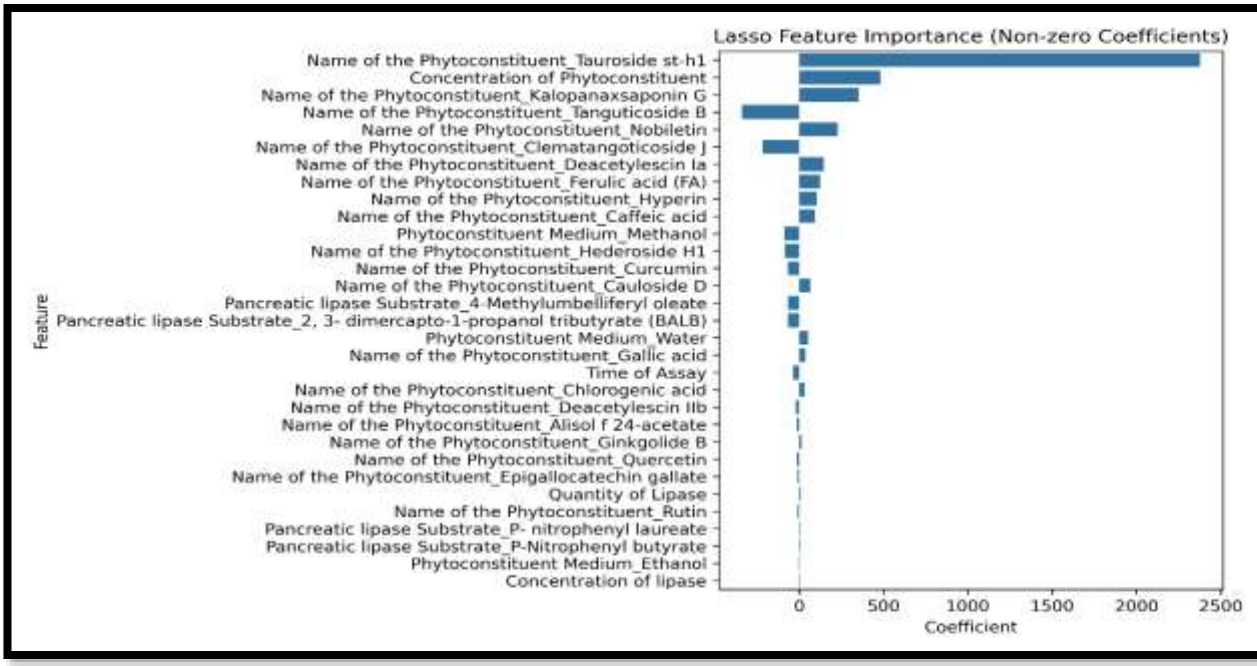

Fig. 3: Non Zero Coefficients- Lasso Feature Importance

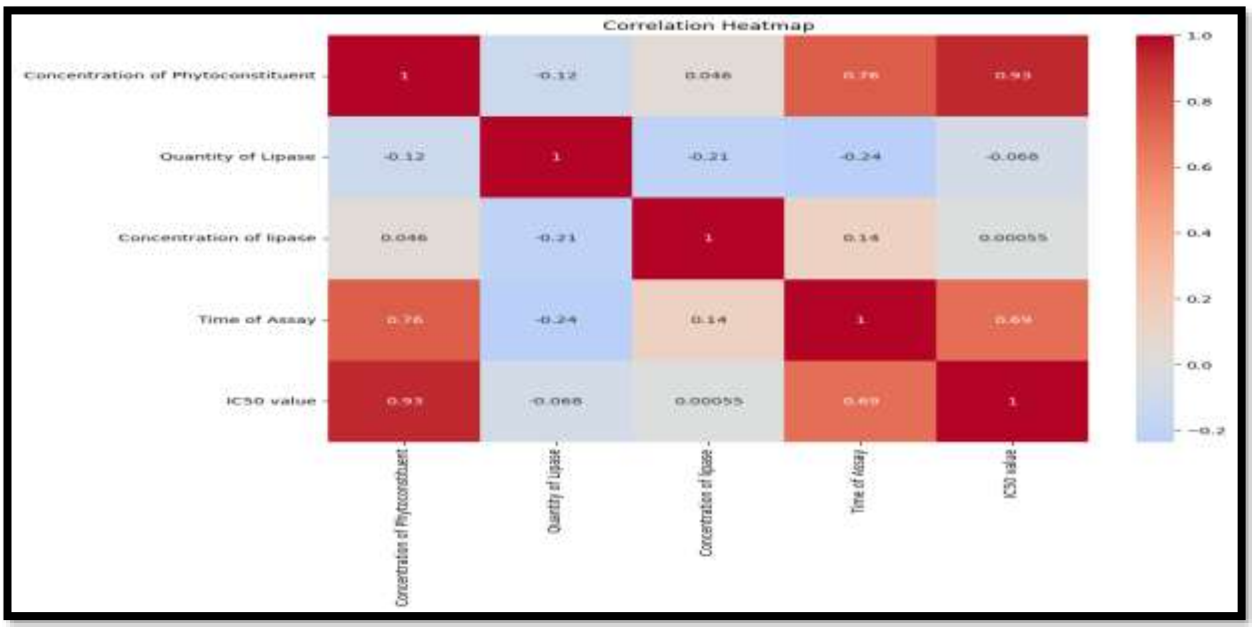

Fig.4: Correlation Heat Map Analysis

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 5,2025
https://theaspd.com/index.php

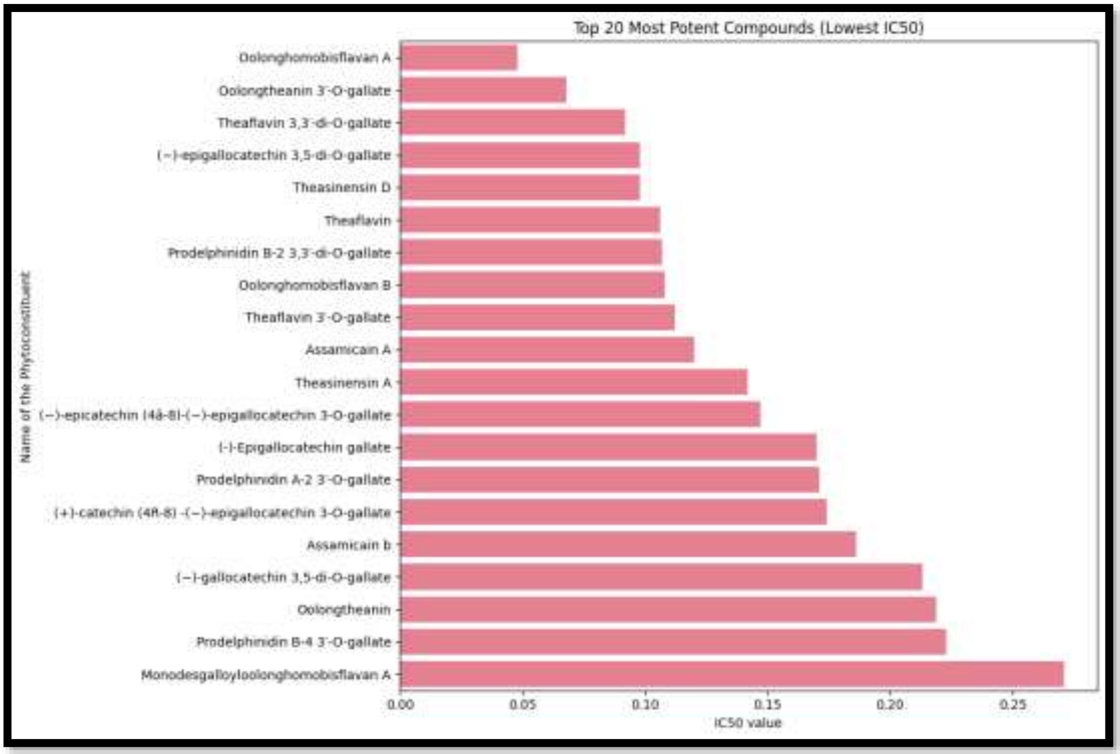

Fig.5: Top 20 most Potent Compounds

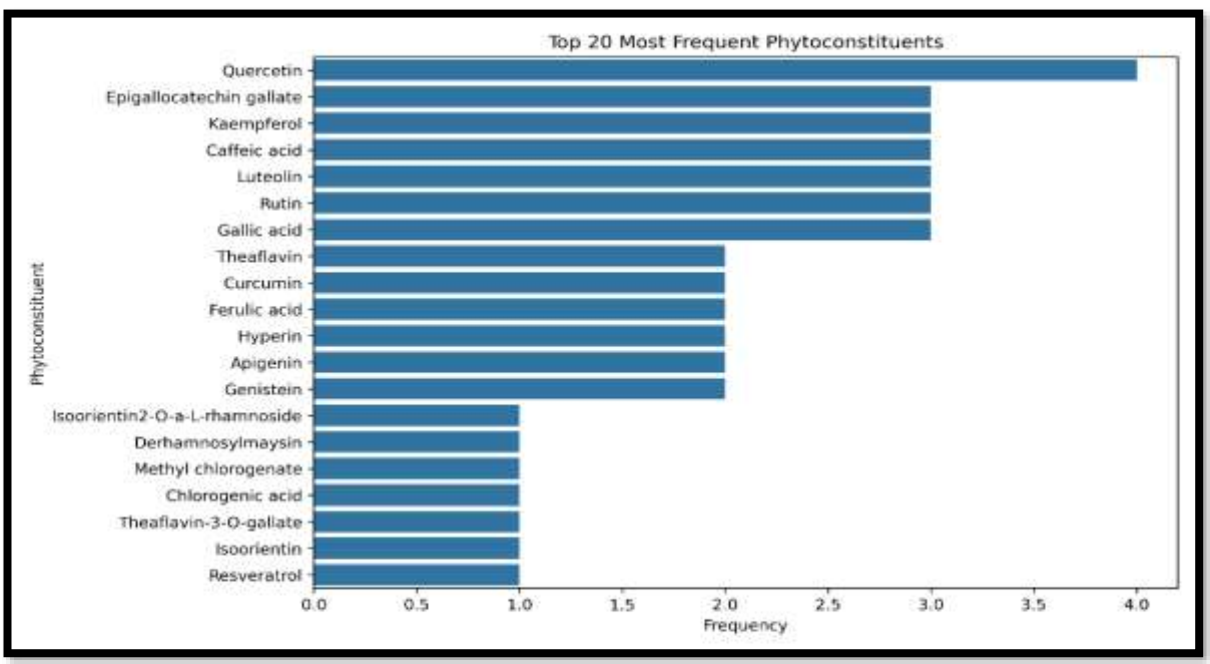

Fig.6: Top 20 Most Frequent Phytoconstituents

REFERENCES

1. Almasri IM (2020) Computational approaches for the discovery of natural pancreatic lipase inhibitors as antiobesity agents. *Future Med Chem* 12(8): 741–757. https://doi.org/10.4155/fmc-2019-0284
2. Alzubi J, Nayyar A, Kumar A (2018) Machine learning from theory to algorithms: An overview. *J Phys Conf Ser* 1142: 012012. https://doi.org/10.1088/1742-6596/1142/1/012012
3. Bai Q, Tan S, Xu T, Liu H, Huang J, Yao X (2021) MolAICal: A soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. *Brief Bioinform* 22(3): bbaa161. https://doi.org/10.1093/bib/bbaa161
4. Balasundaram P, Daley SF (2025) Public health considerations regarding obesity. *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK572122/
5. Bustos AS, Håkansson A, Linares-Pastén JA, Peñarrieta JM, Nilsson L (2020) Interaction of quercetin and epigallocatechin gallate (EGCG) aggregates with pancreatic lipase under simplified intestinal conditions. *PLoS One* 15(4): e0224853. https://doi.org/10.1371/journal.pone.0224853
6. Chan Y, Ng SW, Tan JZX, Gupta G, Negi P, Thangavelu L, Balusamy SR, Perumalsamy H, Yap WH, Singh SK, Caruso V, Dua K, Chellappan DK (2021) Natural products in the management of obesity: Fundamental mechanisms and pharmacotherapy. *S Afr J Bot* 143: 176–197. https://doi.org/10.1016/j.sajb.2021.07.035
7. Chen SY, Zhang QF, Shen HS, Lin SD (2025) Metabolic syndrome prevention potential of tamarillo: Phytochemical composition, antioxidant activity, and enzyme inhibition before and after digestion. *Foods* 14(7): 1282. https://doi.org/10.3390/foods14071282
8. Choudhari S, Patil SK, Rathod S (2023) Identification of hits as anti-obesity agents against human pancreatic lipase via docking, drug-likeness, in-silico ADME(T), pharmacophore, DFT, molecular dynamics, and MM/PB(GB)SA analysis. *J Biomol Struct Dyn* 42(20): 10688–10710. https://doi.org/10.1080/07391102.2023.2258407
9. Destina Ekingen Genc, Ozlem Ozbek, Burcu Oral, Yıldırım R, Ileri Ercan N (2024) Inhibitory potential of selected compounds against pancreatic lipase: A computational and in vitro approach. *ACS Omega* 9(1): 413–421. https://doi.org/10.1021/acsomega.3c05861
10. Emerenciano VP, Barbosa KO, Scotti MT, Ferreira MJP (2007) Self-organizing maps in chemotaxonomic studies of Asteraceae: A classification of tribes using flavonoid data. *J Braz Chem Soc* 18(5): 891–899. https://doi.org/10.1590/S0103-50532007000500004
11. Gao HX, Liang HY, Chen N, Shi B, Zeng WC (2022) Potential of phenolic compounds in *Ligustrum robustum* (Rxob.) Blume as antioxidant and lipase inhibitors: Multi-spectroscopic methods and molecular docking. *J Food Sci* 87(2): 651–663. https://doi.org/10.1111/1750-3841.16020
12. He XQ, Zou HD, Liu Y, Chen XJ, Atanasov AG, Wang XL, Xia Y, Ng SB, Matin M, Wu DT, Liu HY, Gan RY (2024) Discovery of curcuminoids as pancreatic lipase inhibitors from medicine-and-food homology plants. *Nutrients* 16(15): 2566. https://doi.org/10.3390/nu16152566
13. Huang H, Han MH, Gu Q, Wang JD, Zhao H, Zhai BW, Nie SM, Liu ZG, Fu YJ (2023) Identification of pancreatic lipase inhibitors from *Eucommia ulmoides* tea by affinity-ultrafiltration combined UPLC-Orbitrap MS and in vitro validation. *Food Chem* 426: 136630. https://doi.org/10.1016/j.foodchem.2023.136630
14. Jing Y, Luo L, Zeng Z, Zhao X, Huang R, Song C, Chen G, Wei S, Yang H, Tang Y, Jin S (2024) Targeted screening of curcumin derivatives as pancreatic lipase inhibitors using computer-aided drug design. *ACS Omega* 9(25): 27669–27679. https://doi.org/10.1021/acsomega.4c03596
15. Kottekad S, Dandamudi U (2025) Phytochemical analysis of *Triphala* extract, in vitro and in silico evaluation of pancreatic lipase inhibition for obesity management. *Plant Foods Hum Nutr* 80(1): 65. https://doi.org/10.1007/s11130-025-01303-0
16. Kumar A, Chauhan S (2021) Pancreatic lipase inhibitors: The road voyaged and successes. *Life Sci* 271: 119115. https://doi.org/10.1016/j.lfs.2021.119115
17. Liu Y, Pan F, Wang O, Zhu Z, Li Q, Yang Z, Tian W, Zhao L, Zhao L (2023) QSAR model of pancreatic lipase inhibition by phenolic acids and their derivatives based on machine learning and multi-descriptor strategy. *J Agric Food Res* 14: 100783. https://doi.org/10.1016/j.jafr.2023.100783
18. Lu WW, Chen X, Ni JL, Cai WJ, Zhu SL, Fei AH, Wang XS (2021) Study on the medication rule of traditional Chinese medicine in the treatment of acute pancreatitis based on machine learning technology. *Ann Palliat Med* 10(10): 10616–10625. https://apm.amegroups.org/article/view/82028/html
19. Modanwal S, Maurya AK, Mishra SK, Mishra N (2023) Development of QSAR model using machine learning and molecular docking study of polyphenol derivatives against obesity as pancreatic lipase inhibitor. *J Biomol Struct Dyn* 41(14): 6569–6580. https://doi.org/10.1080/07391102.2022.2109753
20. Naquvi KJ (2024) Natural polyphenol-rich inhibitors of pancreatic lipase for obesity management – A systematic review. *J Angiother* 8(10): 1–17. https://publishing.emanresearch.org/Journal/abstract/angiotherapy-8109866
21. Prieto-Rodríguez JA, Lévuok-Mena KP, Cardozo-Muñoz JC, Parra-Amin JE, Lopez-Vallejo F, Cuca-Suárez LE, Patiño-Ladino OJ (2022) In vitro and in silico study of the α-glucosidase and lipase inhibitory activities of chemical constituents from *Piper cumanense* (Piperaceae) and synthetic analogs. *Plants* 11(17): 2188. https://doi.org/10.3390/plants11172188
22. Rocha S, Proença C, Araújo AN, Freitas M, Rufino I, Aniceto N, Silva AMS, Carvalho F, Guedes RC, Fernandes E (2025) Flavonoids as potential modulators of pancreatic lipase catalytic activity. *Pharmaceutics* 17(2): 163. https://doi.org/10.3390/pharmaceutics17020163

23. Sarmah N, Mehtab V, Bugata LSP, Tardio J, Bhargava S, Parthasarathy R, Chenna S (2022) Machine learning aided experimental approach for evaluating the growth kinetics of *Candida antarctica* for lipase production. *Bioresour Technol* 352: 127087. https://doi.org/10.1016/j.biortech.2022.127087
24. Slanc P, Doljak B, Kreft S, Lunder M, Janes D, Strukelj B (2009) Screening of selected food and medicinal plant extracts for pancreatic lipase inhibition. *Phytother Res* 23(6): 874–877. https://doi.org/10.1002/ptr.2718
25. Sood S, Mittal N, Singh TG, Devi S (2023) Pathogenesis of obesity-associated cardiovascular diseases: Key role of biomolecules. *Health Sci Rev* 7: 100098. https://doi.org/10.1016/j.hsr.2023.100098
26. Wójcikowski M, Ballester PJ, Siedlecki P (2017) Performance of machine-learning scoring functions in structure-based virtual screening. *Sci Rep* 7: 46710. https://doi.org/10.1038/srep46710
27. Zhang J, Pan QS, Qian XK, Zhou XL, Wang YJ, He RJ, Yang L (2022) Discovery of triterpenoids as potent dual inhibitors of pancreatic lipase and human carboxylesterase 1. *J Enzyme Inhib Med Chem* 37(1): 629–640. https://doi.org/10.1080/14756366.2022.202985
28. Zhai I, Wang K, Yu Z, Zhou S, Fan J (2025) Pancreatic lipase inhibitors: Virtual screening and mechanistic analysis. *Int J Biol Macromol* 310(Pt 2): 143128. https://doi.org/10.1016/j.ijbiomac.2025.143128