Screening of β2-Receptor Inhibitors with Machine Learning

Mehmet Ali YUCEL[a], Oztekin ALGUL[a,b]

[a]*Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Mersin University, Mersin, Turkey*
[b]*Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Erzincan Binali Yildirim University, Erzincan, Turkey*

mayyucel97@gmail.com

Abstract

Machine learning is increasing in use at various stages of drug development. Identifying the structure-activity relationships of molecules is a substantial step in drug development. β2-adrenergic receptors are associated with common diseases such as hypertension and asthma in society. In this study, a machine learning model that learns the structure-activity relationship for β2-adrenergic receptor inhibitors has been established. For this, initially, the molecule sets were obtained from ChEMBL then the model was trained with SVM and XGBClassifier which are major machine learning algorithms. Finally, the model was evaluated with various metrics.

***Keywords*:** Machine learning, drug discovery, β2-adrenergic receptors, artificial intelligence

Introduction

Artificial intelligence (AI) aims to automate the intellectual actions that humans can perform by machines. Machine learning is a subfield of AI. Classic AI methods can succeed in tasks with specific rules, but not sufficiently successful in complex tasks such as image classification, speech recognition, language translation. Machine learning algorithms achieve more successful results in complex tasks than classic AI. (1)

One of the reasons machine learning has become very popular is that a lot of data has been produced in every field today and it is still being produced. Machine learning is an effective way of turning this data into knowledge. Since drug research and development studies are essential in terms of people's quality of life, the number of data obtained in this field is high. For example, ChEMBL, a molecular database, has more than 2 million different bioactive molecules. (2)

The success of ML methods in computer vision and natural language processing has also shed light on drug development. One of the first notable examples is the success of deep neural networks in the Kaggle competition held by Merck in 2012. Another example study in 2019, discovering Potent inhibitors for "discoidin domain receptor 1 (DDR1)" by In silico Medicine researchers in as little as 21 days. ML can be used in distinct stages of drug development such as the estimation of the chemical properties of the molecule and toxicity profiles. With ML, more rational drug design becomes achievable, and as a result, molecules with high activity, low toxicity, and fewer side effects profiles can be obtained in less time, at less cost. (3)

In this study, an artificial intelligence model was built that can predict the activity of molecules via understanding the relationship between molecules and their targets by analyzing the molecule sets of biological targets. β-adrenoreceptors have been selected as the biological target. Its widespread presence in the body and its diverse functions were a crucial criterion in its selection as a target. For example, the main clinical use of β2-agonists is asthma treatment. With the activation of β2-adrenoreceptors, relaxation occurs in the flat muscles of the bronchi and the airways expand. On the other hand, the main use of β-antagonists in the treatment of angina and hypertension. Drugs that block β1-receptors in the heart reduce heart rate and convulsion. In addition, β blockers contribute to the decrease in blood pressure with their effects in other parts of the body. (4)

Results and Discussion

Material and Method: Support vector machines (SVM) is a supervised machine learning algorithm that can be used in classification analysis. SVM aims to create the hyper plane (decision surface) that best separates classes. In this sense, it aims to maximize classification margins so that the new data is on the appropriate side of the plane. (5)

SVM tries to classify it as linear by maximizing the distance between the hyper plane and the closest instances of the classes (support vector). SVM does not have to make the separation linearly, it can be selected as a polynomal, radial-based or sigmoid core function in accordance with the data set. Choosing the appropriate core functionality is a critical process because different core functions perform differently in the learning process. Although the complexity

of the data is an important criterion for selection, there is no core function specified for each data type. (5,6)

C and gamma are significant parameters of SVM. Parameter C relates to the penalty for incorrect predictions, and the different values of this parameter can affect each support vector based on its associated error. In this sense, small C values indicate small penalty or large margins (distance) for support vectors. The gamma parameter from nonlinear kernels relates to the variance and bias of variables; small gamma values can lead to large variance and low deviation models, and vice versa. (7)

In this study, SVM was used as a machine learning algorithm. Data sets are prepared before the modeling process is advanced. The data is from MuSSeL, which generates using ChEMBL's data. Molecules with activity results for the β2-adrenergic receptors in the data set have various properties such as SMILES codes, chembl numbers, activity results (IC_{50}, nM), hydrogen bond acceptors, logP.

Another algorithm used in this study is XGBoost. XGBoost (Extreme Gradient Boost) is a machine learning technique for regression and classification problems based on the Gradient Boosting Decision Tree (GBDT). Optimizes the traditional GBDT algorithm for the loss function, editing, and segmentation point search algorithm. One of the essential parameters in XGBoost models is "max_depth" which represents the depth of each tree. In general, if the value of this parameter is too large, the training set error in the iteration will decrease rapidly, causing overfitting in the model. (6,8)

Converted to pIC50 values to reduce the variance of activity results in the data set. The threshold value was set at pIC50 7.5; Molecules above 7.5 were active, while that underneath was labeled as inactive. (9) Fingerprints of molecules were created by ECFP6 method. (10) It was then separated into a data set, test, and training set (0.2 and 0.8 respectively). Then SVM models and XGBoost classifier model were created with various parameters. K-fold cross-validation (k=10) was used when creating models, and the results show averages. The results of the models by various metrics are shown in the following table. Python's rdkit, sklearn,

numpy, and pandas libraries were used to perform these operations. Models have trained through Google Colaboratory.

Model	Accuracy score	Matthew Korelasyon sabiti	Recall Score	Precision Score
SVM (kernel=rbf, C=10)	0.87	0.75	0.82	0.89
SVM (kernel=poly, C=10)	0.87	0.74	0.81	0.90
SVM (kernel=rbf, C=1)	0.88	0.77	0.84	0.90
SVM (kernel=poly, C=1)	0.86	0.73	0.77	0.91
XGBClassifier (max_depth=10, learning_rate=0.1)	0.88	0.76	0.86	0.87

Acknowledgement: This research work is financially supported by the Mersin University Research Fund (Project No: 2021-2-TP2-4518).

References

1) Yang X, Wang Y, Byrne R, Schneider G, Yang S. *Chem Rev.* **2019**, 119, 10520.
2) Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. *Drug Discov. Today* **2018**, 23, 1241.
3) Deng J, Yang Z, Ojima I, Samaras D, Wang F. Brief Bioinform. **2022,** 23, 1.
4) Vilar Santiago, Sobarzo-Sanchez Eduardo, Santana Lourdes and Uriarte Eugenio, Curr. Med. Chem. **2017**, 24, 4340.
5) Heikamp K, Bajorath J. *Expert Opin Drug Discov*. **2014,** 9, 93.
6) Zhenqin Wu, Bharath Ramsundar, Evan N. Feinberg, Joseph Gomes, Caleb Geniesse, Aneesh S. Pappu, Karl Leswing, Vijay Pande, arXiv: **2017,** 1703.00564,
7) Vinicius Gonçalves Maltarollo, Thales Kronenberger, Gabriel Zarzana Espinoza, Patricia Rufino Oliveira & Kathia Maria Honorio, *Expert Opin Drug Discov.,* **2018**, 14, 23.
8) Xing G, Liang L, Deng C, Hua Y, Chen X, Yang Y, Liu H, Lu T, Chen Y, Zhang Y. *ACS Comb Sci.* **2020**,22, 873.
9) Lane T, Russo DP, Zorn KM, Clark AM, Korotcov A, Tkachenko V, Reynolds RC, Perryman AL, Freundlich JS, Ekins S. *Mol Pharm*. **2018**, 15, :4346.
10) Rogers D, Hahn M. *J Chem Inf Model*. **2010** May 24; 50,742. doi: 10.1093/bib/bbab430.

10. INTERNATIONAL DRUG CHEMISTRY CONFERENCE

MARCH 10-13, 2022 KEMER / ANTALYA

THE CHEMISTS SOCIETY

CONGRESS PROGRAM

Ana Sponsor

sem

Sponsorlar

ATABAY KİMYA SANAYİ ve TİCARET A.Ş.
redoks LAB
ANT TEKNİK
VEFA İLAÇ
AYES
LiKrom
AN-KA
info
BMS KİMYA
bilim
SPRINGER NATURE
WMIN LAB
GVS
WHITECHNIC
DORUK ANALİTİK
ChromaScience Chemistry Technologies Ltd.Şti.
ULKAR KİMYA SANAYİİ VE TİCARET A.Ş.
YGT LABORATUVAR
HEKA SCIENCE
İKMİB İstanbul Kimyevi Maddeler ve Mamulleri İhracatçıları Birliği
SUDUM

Destekleyenler

SURDER SAĞLIK ÜRÜNLERİ DERNEĞİ
LabMedya
MEDICHEM
MonoLab Laboratuvar Çözümleri
labakademi
SasuChem
Made in Türkiye
Pharma Türkiye
Beauty Türkiye
Prinova NAGASE Group

http://www.ilackongresi.org/

CHAIR

Prof. Dr. Öztekin **ALGÜL**

VICE CHAIRS

Assoc. Prof. Serdar **BURMAOĞLU**

SCIENTIFIC COORDINATORS

Prof. Dr. Aliye **ALTUNDAŞ**

Prof. Dr. Hasan **SEÇEN**

Prof. Dr. Mehlika Dilek **ALTINTOP**

Prof. Dr. Nermin **ŞİMŞEK KUŞ**

Prof. Dr. Özdemir **DOĞAN**

Assoc. Prof. Esen Bellur **ATİCİ**

CONFERENCE SECRETARATS

Assist. Prof. Derya **AKTAŞ ANIL**

Assist. Prof. M. Fatih **POLAT**

CONFERENCE COORDINATOR

İkram **CENGİZ,** Turkish Chemist Society

ADVİSORY BOARD

Prof. Dr. Erden **BANOĞLU**

Prof. Dr. Metin **BALCI**

Prof. Dr. Ramazan **ALTUNDAŞ**

Dr. O. Mutlu **TOPAL**

Dr. Y. Lütfü **ŞEN**

Mahmut **ÖZBEK**

ORGANIZATION COMMITTEE

Prof. Dr. Atilla **AKDEMİR,** Bezmialem Vakıf University

Prof. Dr. Ahmet Ceyhan **GÖREN**, Gebze Technical University

Prof. Dr. Yunus **KARA**, Atatürk University

Prof. Dr. Cumhur **KIRILMIŞ**, Adıyaman University

Prof. Dr. Serdar **ÜNLÜ**, Ali Raif

Prof. Dr. Rahime **ŞİMŞEK**, Hacettepe University

Assoc. Prof. Ufuk **ATMACA**, Atatürk University

Assoc. Prof. Kayhan **BOLELLİ**, Ankara University

Assoc. Prof. Yasin **ÇETİNKAYA**, Atatürk University

Assoc. Prof. Serdar **KARAKURT**, Selçuk University

Assoc. Prof. Ömer Faruk **KARATAŞ**, Erzurum Technical University

Assoc. Prof. Haydar **KILIÇ**, Atatürk University

Assoc. Prof. Nurettin **MENGEŞ**, Van Yüzüncü Yıl University

Assoc. Prof. Sema **MISIR**, Cumhuriyet University

Assoc. Prof. Bilal **NİŞANCI**, Atatürk University

Assoc. Prof. Salih **ÖZÇUBUKÇU**, ODTÜ

Assoc. Prof. Belgin **SEVER**, Anadolu University

Assoc. Prof. Elif **ŞENKUYTU**, Atatürk University

Assoc. Prof. Serhan **TÜRKYILMAZ**, ODTÜ

Dr. Levent **KANDEMİR,** Polifarma

Dr. Hasan **KOYUNCU**, Ulkar Kimya

Dr. Yusuf **ÖZLÜ**, Mustafa Nevzat İlaç

Dr. İrem **YENİCE**, Gensenta İlaç

Dr. Emine **YURTOĞLU**, Hektaş

Dr. Mustafa **ADIYAMAN**, SANOFİ

Dr. Büşra **ÖZTÜRK AYDIN**, Atatürk University

Dr. Ümit **DERELİ**, AİFD

Dr. Ömer Faruk **FIRAT**, Polifarma

Fatma Gözde **AYDIN**, DEVA

Arzu **GÖBEK**, Atatürk University

Figen Onuk **GÖREN**, Pharmactive

Adil **KAYA**, Atabay İlaç Sanayi

Gürmen **KAYNAR**, Polifarma

Mustafa **KÖKTÜRK**, Biofarma İlaç

Melda **MISIRLI**, Sanovel/Arven

Melahat **ÖZARAÇ**, Santa Farma

Övünç **SARIKAYA**, Novartis

Ali İhsan **SEÇKİN**, Arven

Sercan **SEMİZ**, World Medicine

Müge **SÜRMEN**, Abdi İbrahim

Ömer **SÜRMEN**

Özgür **YILMAZ**, TÜBİTAK

Koray **TANRIVER**, Abdi İbrahim

Nazım **TÜRKAY**, Koçak Farma

Ersin **YILDIRIM**, Santa Farma

SCIENTIFIC COMMITTEE

Prof. Dr. Özlen GÜZEL **AKDEMİR**, İstanbul University

Prof. Dr. Burhan **ATEŞ**, İnönü University

Prof. Dr. Mutlu **AYTEMİR**, İzmir Katip Çelebi University

Prof. Dr. Mustafa **CEYLAN**, GOP University

Prof. Dr. Phillip **BOWEN**, Mercer University, ABD

Prof. Dr. Yahya **GÜZEL**, Erciyes University

Prof. Dr. Ali **DİŞLİ**, Gazi University

Prof. Dr. Masami **OTSUKA**, R&D Science Farm Ltd, Japan

Prof. Dr. Adil **MARDİNOĞLU**, King's College, UK

Prof. Dr. Claus **SCHNEIDER**, Vanderbilt University, ABD

Prof. Dr. Ahmet **ÖZDEMİR**, Anadolu University

Prof. Dr. Aykut **ÜREN**, Georgetown University, ABD

Prof. Dr. Hasan **TÜRKEZ**, Atatürk University

Prof. Dr. İlkay **YILDIZ**, Ankara University

Prof. Dr. Ferhan **TÜMER**, Kahramanmaraş Sütçü İmam University

Assoc. Prof. Hayrettin Ozan **GÜLCAN**, Doğu Akdeniz University

Assist. Prof. Mohammad A. **OBEID**, Yarmouk University, Jordan

Assist. Prof. Simone **CARRADORİ**, Università degli Studi G. d'Annunzio Chieti e Pescara, Italy

Assist. Prof. Antoine **MARION**, ODTÜ

Assist. Prof. Hacer **KARATAŞ BRISTOW**, İstanbul Medipol University

Dr. Gökçe **YILDIRIM BUHARALIOĞLU**, Ege University

Dr. Halilibrahim **CİFTCİ**, R&D. Science Farm Ltd, Japan

Dr. Tuğba **BOLELLİ**, Ankara University

Dr. Ronak Haj **ERSAN**, Cihan University Duhok, Iraq

Dr. Mikako **FUJİTA**, Kumamoto University, Japan

SOCIAL COMMITTEE

Dr. Tayfun **OLTULU**, NEPHA Tıbbi Cihazlar

Birsen **CELAYİR**, Sanofi

INVITED SPEAKERS

Prof. Dr. Masami **OTSUKA,** Co-founder & R&D Director, Science Farm Ltd.

Prof. Dr. Mikako **FUJITA**, Kumamoto University

Assoc. Prof. Halilibrahim **ÇİFTÇİ**, Co-founder & Vice President, Science Farm Ltd