Computers in Biology and Medicine 164 (2023) 107314

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

ALPACA: A machine Learning Platform for Affinity and selectivity profiling of CAnnabinoids receptors modulators

Pietro Delre [a], Marialessandra Contino [b], Domenico Alberga [a], Michele Saviano [c], Nicola Corriero [a,**], Giuseppe Felice Mangiatordi [a,*]

[a] CNR – Institute of Crystallography, Via Amendola 122/o, 70126, Bari, Italy
[b] Department of Pharmacy - Pharmaceutical Sciences, University of Bari "Aldo Moro", via E. Orabona, 4, I-70125, Bari, Italy
[c] CNR – Institute of Crystallography, Via Vivaldi 43, 81100, Caserta, Italy

ARTICLE INFO

Keywords:
Machine learning
Ligand-based models
Classifiers
Web-platform
Cannabinoid receptors

ABSTRACT

The development of small molecules that selectively target the cannabinoid receptor subtype 2 (CB2R) is emerging as an intriguing therapeutic strategy to treat neurodegeneration, as well as to contrast the onset and progression of cancer. In this context, in-silico tools able to predict CB2R affinity and selectivity with respect to the subtype 1 (CB1R), whose modulation is responsible for undesired psychotropic effects, are highly desirable. In this work, we developed a series of machine learning classifiers trained on high-quality bioactivity data of small molecules acting on CB2R and/or CB1R extracted from ChEMBL v30. Our classifiers showed strong predictive power in accurately determining CB2R affinity, CB1R affinity, and CB2R/CB1R selectivity. Among the built models, those obtained using random forest as algorithm proved to be the top-performing ones (AUC in validation $\geq$0.96) and were made freely accessible through a user-friendly web platform developed *ad hoc* and called ALPACA (https://www.ba.ic.cnr.it/softwareic/alpaca/). Due to its user-friendly interface and robust predictive power, ALPACA can be a valuable tool in saving both time and resources involved in the design of selective CB2R modulators.

1. Introduction

Cannabinoid receptor 1 (CBR1) and cannabinoid receptor 2 (CB2R) are receptor subtypes belonging to the EndoCannabinoid System (ECS) [1,2], responsible for the network of signalling ECS-mediated involved in several disorders based on an inflammatory state, such as neurodegenerative diseases, cancer and neuropathic pain [3–5]. CB1R and CB2R are both G-protein-coupled receptors (GPCRs (G_i)) whose activation leads to the inhibition of the adenylate-cyclase enzyme and physiologically differently located: CB1R is mainly present in the Central Nervous System (CNS) while CB2R at the peripheral level [1,2]. Noteworthy, CB2R is overexpressed in inflammatory state, suggesting that developing selective CB2R agents may be an effective strategy for treating neuroinflammation [6,7]. In particular, there are pieces of evidence that the overexpression of CB2R in activated microglia and its activation may induce a final anti-inflammatory activity and, thus, neuroprotection without the psychotropic side effects typical of the central CB1R stimulation [7,8]. Furthermore, in other diseases where inflammation is the cause of the onset as cancer, covid-19, obesity or diabetes, the development of selective CB2R ligands may have a significant impact [9,10]. Thus, the need of developing CB2R selective agents to treat/combat inflammation is a challenging aim considering the utility to monitoring (diagnostic field) the disease progression detectable by the CB2R changed expression and to potentially treat the above-mentioned diseases (therapeutic field). This is a challenging issue since the two receptors share a high sequence homology (44%) and it is not so easy to hit CB2R without modulating also the other subtype [11]. This increasing interest in targeting the CB2R for various pathological conditions has prompted numerous research groups to dedicate their efforts to designing CB2R ligands. Remarkably, a great impetus in this direction has been provided by the recent deposition of the first x-ray structures of human CB2R in the Protein Data Bank (PDB IDs: 5ZTY [12] and 6KPC [13]) hence providing the structural information required for using structure-based approaches such as for instance molecular docking.

* Corresponding author.
** Corresponding author.
E-mail addresses: nicola.corriero@ic.cnr.it (N. Corriero), giuseppe.mangiatordi@ic.cnr.it (G.F. Mangiatordi).

https://doi.org/10.1016/j.compbiomed.2023.107314
Received 13 June 2023; Received in revised form 10 July 2023; Accepted 7 August 2023
Available online 7 August 2023

Noteworthy, using these X-ray structures, we recently developed active agents as single target ligands [14] and MultiTarget Directed Ligands (MTDLs), driven by the importance of this class of compounds (MTDLs) in multifactorial disorders as the neurodegenerative ones [14–16]. Furthermore, Jing-Fang Yang et al. used molecular docking and molecular dynamics simulations to determine the binding poses of long-chain endocannabinoids within the CB1R and CB2R binding site, demonstrating the usefulness of structure-based strategies to design more effective long-chain endocannabinoids [17]. However, despite these achievements, reaching CB2R selectivity using standard structure-based approaches is still considered challenging due to the high similarity between the CBRs binding sites, so that alternative strategies are highly desirable such as advanced approaches based on MD simulations and binding free energy calculations, recently proved to be useful for reaching selectivity in the presence of similar binding sites [18–22]. Noteworthy, despite the increasing availability of extensive bioactivity data concerning CB2R and CB2R ligands in freely accessible repositories such as ChEMBL [23] and PubChem [24], only a few efforts have been dedicated to developing ligand-based models for assessing the affinity and selectivity of candidate compounds towards CBRs. To the best of our knowledge, the only existing examples are the recent studies conducted by Bian et *al.*, Ruano-Ordás et *al.,* Zhou et *al.*, and Mizera et al. reporting machine learning (ML) models for predicting CB1R or CB2R activity [25–28]. Nevertheless, these classifiers are not straightforwardly useable by medicinal chemists interested in design CBRs ligands as they were not made available by the authors. The presented study aims to bridge this gap by developing and making freely accessible a novel computational workflow that effectively addresses selectivity issues about CB2R. Leveraging the bioactivity data from ChEMBL, we trained ML models using the atom pair (AP) fingerprint to accurately predict the affinity and selectivity profiles of CB2R ligands. More specifically, the models were trained to *i)* identify molecules with high affinity towards CB2R ($pK_i > 6.5$ or $pK_i > 7$) while excluding affinity towards CB1R ($pK_i > 5.5$ or $pK_i > 6$) and ii) develop a specific classifier for CB2R selectivity. We employed five different classification algorithms to construct our models, which were subsequently tested and validated against two external datasets extracted from independent sources. Our models exhibited excellent performance levels, as demonstrated by various widely accepted metrics and we made the top performing ones freely accessible through a user-friendly web platform developed *ad hoc* and called ALPACA (https://www.ba.ic.cnr.it/softwareic/alpaca/). ALPACA does not require any specific cheminformatics or programming skill, hence it can be considered a valuable resource for medicinal chemists seeking to design new selective CB2R modulators for cancer and neurodegeneration. Furthermore, ALPACA offers the capability to conduct virtual screening campaigns by inputting a compound library, enabling the identification of highly probable CB2R selective molecules with a high level of confidence. With its robust predictive power and user-friendly interface, we believe that ALPACA can serve as an invaluable tool to aid medicinal chemists in the design of selective CB2R modulators for cancer and neurodegeneration. Its ability to conduct virtual screening campaigns further enhances its usability and practicality.

2. Materials and Methods

2.1. Data set preparation

6031 entries, annotated exclusively with K_i measures, were retrieved from the ChEMBL [23] v30 database according to the Target ID (ChEMBL253) assigned to the CB2R channel. To ensure data validity and following an approach successfully employed for other classifiers [29,30], the database was mined by retaining only those entries matching specific criteria: *i*) referring to assays conducted on human targets ("target_organism" = "*Homo sapiens*"), *ii*) marked as direct binding ("assay_type" = "B"), and *iii*) without any warning in the 'data_validity_comment' field (**5369** entries). Although associated with the CHEMBL253 target, **26** entries were eliminated as reporting in the assay description field the following sentence: "Displacement of [H3]-CP-55940 from human recombinant CB1R express in CHO cells" (**11**) and "Displacement of [3H]-SR141716A from human CB1R expressed in CHO membranes after 1 h by liquid scintillation counting" (**15**). SMILES were curated using an in-house automated procedure to remove organometallic and inorganic compounds, chemicals characterized by unusual elements and mixtures, neutralizing salts, and stereochemistry [29,31]. The neutralized SMILES were converted to a standardized QSAR-ready format using OpenBabel [32] implemented in the KNIME Analytics Platform [33] to generate Canonical SMILES. To standardize the data, we converted Ki values from molar concentration (M) to pKi (-log Ki). In the last step, duplicates were aggregated in unique entries and the standard deviation (σ) related to the pK_i values was computed. For compounds (**45**) with a pKi mean σ greater than 2 (outliers), the mean value was excluded from the study. Finally, after removing **1784** entries present as duplicates, the curated dataset consists of **3514** chemicals (hereinafter referred to as *CB2R-qsarDB*) and the corresponding experimental values (pK_i mean). It is worth noting that *CB2R-qsarDB* includes compounds without any recorded pK_i values but noted as 'not active' in the comment field, therefore these have no affinity for CB2R. Using the criteria described above (*i.e.,* the criteria for selecting validated entries from ChEMBL), we retrieved from ChEMBLv30 **6140** entries with Ki measures for CB1R (Target ID: 218) and collected **5504** entries annotated as direct binding and free of warnings. Finally, using the automated procedure, we remove *i)* **1590** duplicate entries and *ii)* **68** entries considered outliers (pK_i mean with a $\sigma > 2$). In doing that, the final database (*CB1R-qsarDB*) contains **3846** entries with the corresponding experimental value (pK_i mean). Compounds denoted as 'not active' in the comment field were considered to have no affinity for CB1R. It is important to highlight that as a result of the rigorous data curation process, a significant majority of the pKi values in the CB2R-qsarDB (78%) and CB1R-qsarDB (82%) datasets were derived from radioligand binding assays. These assays are widely recognized as the gold standard for obtaining reliable and high-quality affinity data. Fig. S1 (available in the Supporting Information), reports the pK_i distribution of all the molecules belonging to *CB2R-qsarDB* and *CB1R-qsarDB* datasets. Notice that "not active" compounds were excluded from this analysis. It is worth to note, moreover, that both *CB2R-qsarDB* and *CB1R-qsarDB* exhibited high values of internal diversity (0.72 and 0.75, respectively) as calculated by the mean Tanimoto distances among each molecule and all others within the same dataset. This demonstrates that both datasets cover a wide range of chemical space. In this work, we selected two different thresholds for predicting CB2R ($pK_i = 6.5$ and 7) and CB1R ($pK_i = 6$ and 5.5) affinity. In other words, we aimed at discriminating molecules with high affinity (HAF) from those with low or absent affinity (LAF). Additionally, we merged *CB1R-qsarDB* and *CB2R-qsarDB* to create a new dataset, comprising those molecules having experimental data for both the receptors and high affinity for at least one of them. In particular, we removed all the compounds with no data for both receptors, as well as those with low affinity based on defined thresholds: $pK_i = 6.5$ for CB2R and $pK_i = 5.5$ for CB1R. In doing that, this new database (*CBRs-qsarSEL*) consists of **2183** compounds. These molecules were labelled based on the pK_i difference computed as follows:

$$pK_{i(CB2R-CB1R)} = pK_{i(CB2R)} - pK_{i(CB1R)} \quad (1)$$

Specifically, we built three classes: i) *no-sel* for molecules with $-1 \leq pK_{i(CB2R-CB1R)} \leq 1$ (non-selective); ii) *CB2R-sel* for those with $pK_{i(CB2R-CB1R)} > 1$, selective towards CB2R, and iii) *CB1R-sel* for the remaining compounds with $pK_{i(CB2R-CB1R)} < -1$. Finally, we represented each SMILES string in the database as an Atom pair fingerprint [34]. In doing that, each chemical substructure was codified by a binary representation (1024 bits) to indicate the presence (1) or absence (0) of

specific characteristics.

2.2. External set preparation

We built two external sets using CB2R and CB1R data stored in Pubchem [24] (pub) and in the binding database (bdb) [35]. To achieve this, we downloaded **6098** (Pubchem) and **13,159** (BindingDB) entries annotated as CB2R Ki measures. For PubChem, we recovered the relative SMILES strings using an in-house Python script with the PubChemPy library, allowing chemical searches by the compound identifier (CID). We then curated these databases using the in-house KNIME Workflow described above to remove compounds already included in *CB2R-qsarDB*. As for the Binding DB is concerned, we also removed those compounds already present in PubChem. As a result, we built two External datasets (ES(pub)/ES(bdb)) consisting of **373** and **355** compounds, respectively. As for CB1R data, we initially collected **6133** entries from PubChem and **14,568** from Binding DB, which became **238** (ES(pub)/**268** (ES(bdb)) after the data curation pathway. We then merged the ES of CB2R and CB1R databases and followed the approach described above for building *ES_CBRs*. As a result, we created a new ES consisting of **129** (selective and non-selective) compounds based on their $pK_{i(CB2R-CB1R)}$ values.

2.3. Data set splitting

Following a rational splitting approach, we split each starting dataset (*CB1R-qsarDB* and *CB2R-qsarDB*) into a training set (TS) and a validation set (VS). Notably, the RDkit Picker Diversity node [36] was employed separately on the two classes (*i.e.*, HAF and LAF) resulting from the considered affinity thresholds ($pK_i = 6.5$ and $pK_i = 7$ for CB2R; $pK_i = 6$ and $pK_i = 5.5$ for CB1R). This node utilizes the considered fingerprint, atom pair (AP) in our case, and picks 80% of the most diverse molecules based on the Tanimoto distance [37]. In doing so, the resulting TS included 80% of the total molecules while the remaining 20% constituted the VS. Noteworthy, several studies underline the remarkable predictive power of these fingerprints in ligand-based models, being proved to be able to successfully capture essential structural details of small organic molecules [38]. We applied the same approach to the three classes of *CBRs-qsarSEL* (*i.e., CB1R-sel* (selective towards CB1R), *CB2R-sel* (towards CB2R), and *no-sel* (non-selective)). Table 1 summarizes, for each case, the final composition of TS, VS, and ES after the applicability domain (AD) evaluation [39,40] (see **section 2.6** for the methodological details) and the imbalanced ratio (IR) calculated as the ratio between the number of majority and minority instances [41]. Note that such a procedure allowed us to keep fixed the ratio between the classes in each subset. Furthermore, we assessed the coverage of the chemical space by the generated TS using a principal component analysis (PCA) [42]. To accomplish this, we calculated 16 physicochemical descriptors for each ligand using the CDK toolkit and standardized them using the Normalizer KNIME node [43]. The PCA simplified the high-dimensional data deriving from the computed descriptors and reduced it to two principal components (PC1 and PC2) that account for 70% of the variance. Then, we plotted each ligand in a 2D chemical space.

2.4. Models implementation and validation

We developed two sets of binary models differing for the considered affinity thresholds ($pK_i = 6.5$ and $pK_i = 7$ for CB2R; $pK_i = 6$ and $pK_i = 5.5$ for CB1R). We used five classification algorithms: Random forest (RF) [44], K-nearest neighbors (KNN) [45], Gradient boosting (GB) [46], eXtreme Gradient Boosting (Chen and Guestrin, 2016) [47], Multi-Layer Perceptron (MLP) [48]. We employed the following KNIME nodes: Tree Ensemble Learner/predictor, K Nearest Neighbor, Gradient Boosted Trees Learner/predictor, and RProp MLP Learner/predictor to develop RF, KNN, GB, and MLP classifiers [33]. As for XGB, we integrated into KNIME an in-house python code employing Pandas and XGboost libraries. Finally, we developed the selectivity model (*i.e.,* built on *CBRs-qsarSEL* data), based on the RF algorithm and using the Tree Ensemble Learner/predictor provided by the KNIME space. In all cases, we found the optimal algorithm setting for training the final model employing a hyperparameter tuning based on 5-fold cross-validation (CV) performance [49]. To do this, we performed a grid search [50], except for XGB and GB, where we used Bayesian optimization to reduce the computational cost [51]. Table S1, available as supplementary material, reports the parameters chosen for each algorithm.

2.5. Metrics

Cooper's statistics, including sensitivity (SE) (**2**), specificity (SP) (**3**), and balanced accuracy (BA) (**4**), were used to evaluate the performance of the classification models [52].

Table 1
Partitioning schemes at each considered affinity threshold. For each dataset employed as training set (TS), validation set (VS) and external set (ES), the total number of chemicals (#), the number of high-affinity (HAF) ligands, the number of low-affinity (LAF) chemicals and the corresponding imbalanced ratio (IR) calculated as the ratio between the number of majority and minority instances are shown.

CB2R	Affinity threshold (pK_i)							
	6.5				7			
	HAF	LAF	IR	#	HAF	LAF	IR	#
TS	1422	1358	1.05	**2780**	1094	1684	1.54	**2778**
VS	377	357	1.06	**734**	320	416	1.30	**736**
ES(pub)	231	142	1.63	**373**	181	192	1.06	**373**
ES(bdb)	165	190	1.15	**355**	141	214	1.52	**355**
CB1R	**5.5**				**6**			
	HAF	LAF	IR	#	HAF	LAF	IR	#
TS	1566	1479	1.06	**3045**	1295	1755	1.36	**3050**
VS	412	389	1.06	**801**	334	462	1.38	**796**
ES(pub)	111	127	1.14	**238**	91	147	1.61	**238**
ES(bdb)	162	106	1.52	**268**	132	136	1.03	**268**
CBRs_sel	**Selectivity**							
	CB2R-sel	*CB1R-sel*	*no-sel*	#				
TS	835	237	674	**1746**				
VS	209	59	169	**437**				
ES	58	21	50	**129**				

$$SE = \frac{TP}{TP + FN} \tag{2}$$

$$SP = \frac{TN}{TN + FP} \tag{3}$$

$$BA = \frac{SE + SP}{2} \tag{4}$$

In equations (3) and (4), True Positives (**TP**) and True Negatives (**TN**) represent the correctly classified positive and negative samples by the models, respectively; false negatives (**FN**) and false positives (**FP**) are the misclassified positive and negative samples, respectively. The MCC (Matthews correlation coefficient) **(5)** was also used to indicate the quality of binary classification, with values ranging from −1 to +1. A value of +1 defines a perfect performance, 0 is random, and −1 represents complete misclassification [53].

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

Finally, we computed the AUC (area under the curve) of the receiver operating characteristic (ROC) to evaluate the predictive accuracy of the developed models. The AUC ranges between 0 and 1, with 0 indicating that all the predictions are incorrect and 1 representing the ideal (optimal) performance (*i.e.,* all predictions are correct). The AUC measures the ability of the models to rank positive compounds in the top positions compared to decoys. In particular, we evaluated the AUC for each model by classifying the compounds according to the estimated probability to be HAF returned by the developed classification model [54].

2.6. Applicability domain

To ensure the reliability of the predictions [39,40] performed in both internal (VS) and external (ES) validation, we defined, for each developed model, the corresponding AD by using the Domain - Similarity node which enables the calculation of the Euclidean distances between the compounds belonging to the TS and those to be predicted [55,56]. In particular, this approach allows defining an AD threshold (ADP) following these steps: *i)* the computation of all the Euclidean distances between all the possible pairs of training compounds based on representative descriptors (AP fingerprint in our case); *ii)* the creation of a set of all the distances lower than this average; *iii)* the calculation of the mean (*d*) and the standard deviation (σ) of this set and *iv)* the final definition of the ADP with this equation:

$$ADP = d + Z\sigma \tag{6}$$

where Z is an empirical cut-off value equal to 0.5 by default [55,56]. In plain terms, based on this approach, predictions can be considered reliable only for those compounds returning a distance, with respect to the nearest neighbor compound of the TS, lower the computed APD threshold.

3. Results and discussion

As mentioned above, we developed 20 classifiers based on affinity data extracted from the ChEMBL database [23] (v30) and employing different ML algorithms, namely RF [44], KNN [45], GB [46], XGB [47], and MLP [48], available in the KNIME Analytics Platform (v. 4.5.2) [33]. In other words, for each affinity threshold used to distinguish HAF from LAF ($pK_i = 6.5$ and $pK_i = 7$ for CB2R; $pK_i = 6$ and $pK_i = 5.5$ for CB1R), five models were developed. For the sake of clarity, Fig. 1 displays the

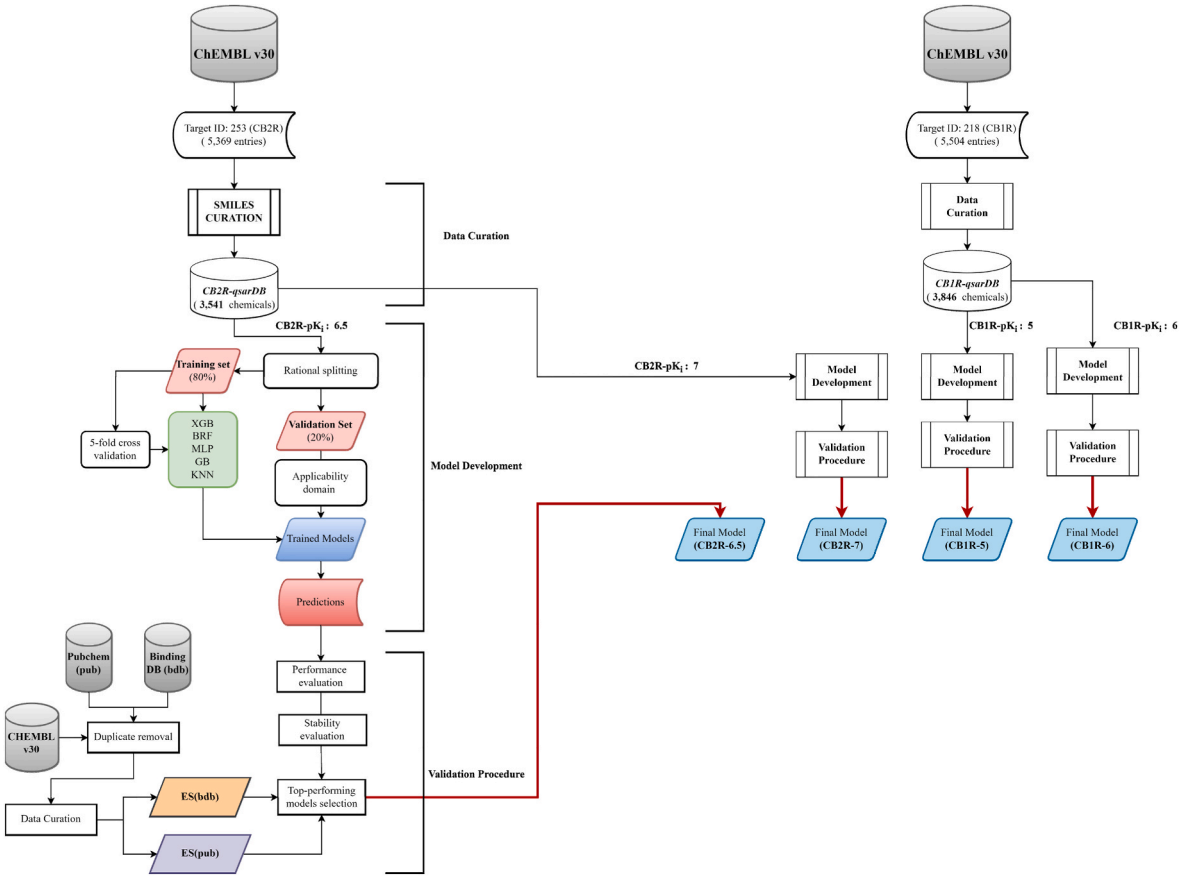

Fig. 1. Flowchart showing the main steps of the adopted computational workflow.

main steps of the adopted computational workflow.

The dataset used to build the models consists of highly curated pK_i values of **3846** (CB1R) and **3514** (CB2R) organic compounds along with the corresponding AtomPair (AP) fingerprints [34]. As depicted in Fig. 2, the datasets employed in this study encompass a diverse range of structural characteristics of CBRs ligands.

It is noteworthy that CB1R and CB2R TS, generated through the rational splitting approach, exhibit a broad chemical space that not only overlaps with the relative VS but also encompasses the two employed ES, which originate from independent resources. This observation implies that the developed models possess the capability to make predictions for compounds belonging to different chemical classes.

3.1. Classification models of CB2R affinity

3.1.1. $pK_i = 6.5$ as affinity threshold

Fig. 3A shows radar plots that compare the performances in VS of the CB2R affinity classifiers built using as threshold $pK_i = 6.5$.

For all the values of the quality metrics obtained from the developed models, we direct the interested reader to Table S2, which is available as supplementary material. It is worth noting that all the classifiers were trained using a balanced dataset, as evident looking at Table 1. In particular, the employed TS and VS returned an IR (ratio between the number of majority and minority instances) [41] equal to 1.05 and 1.06, respectively.

Remarkably, RF, GB, and XGB emerged as the algorithms returning the top-performing models, being responsible for the highest AUC (0.97) and very similar SP and SE values (difference ranging from 0.04 to 0.06). Of note are also the very high MCC (0.80) and BA (0.90) values, strongly supporting the high predictive power of the developed classifiers. The performed 5-fold cross validation (5-CV) further confirms the robustness of the models built using RF, GB and XGB as algorithms, being all responsible for BA $\geq$ 0.78 and AUC $\geq$0.86 (Fig. 4A and Table S3).

Encouraged by these results, these classifiers were challenged on two ES including **373** (*ES(pb)*) and **355** (*ES(bdb)*) compounds extracted from PubChem [24] and Binding DB [35] repositories, respectively. Although it is well known that validating predictive models on compounds coming from external and independent resources is particularly challenging [57], the models returned satisfactory performances, with BA values ranging from 0.76 (XGB) to 0.80 (RF) and from 0.73 (RF) to 0.74 (XGB) on *ES(pub)* (Fig. 5A and Table S4) and *ES(bdb)* (Fig. 6A and Table S4), respectively.

In addition, the robustness and versatility of the models was supported also by the computed AUC values, all $\geq$0.82, irrespective of the considered algorithm and ES.

3.1.2. $pK_i = 7$ as affinity threshold

Satisfactory performances in validation were obtained also when $pK_i = 7$ was used as CB2R affinity threshold (Fig. 3B and Table S2) with, again, RF, GB, and XGB ensuring the highest BA ($\geq$0.90), AUC ($\geq$0.96) and MCC ($\geq$0.81). Noteworthy, herein the use of a higher affinity threshold ($pK_i = 7$ *vs.* $pK_i = 6.5$) makes the TS slight unbalanced (IR $\approx$ 1.5). It is worth mentioning that according to previous studies, models developed using a TS with an IR of 1.5 or less do not show significant improvement from the use of resampling techniques [58]. Therefore, in this study, these approaches were not considered for these models. It is interesting to note that the top-performing models demonstrate comparable values of sensitivity (SE) and specificity (SP), with differences ranging from 0.03 to 0.05. As previously experienced, the 5-fold cross-validation (Fig. 4B and Table S3) and the external validation

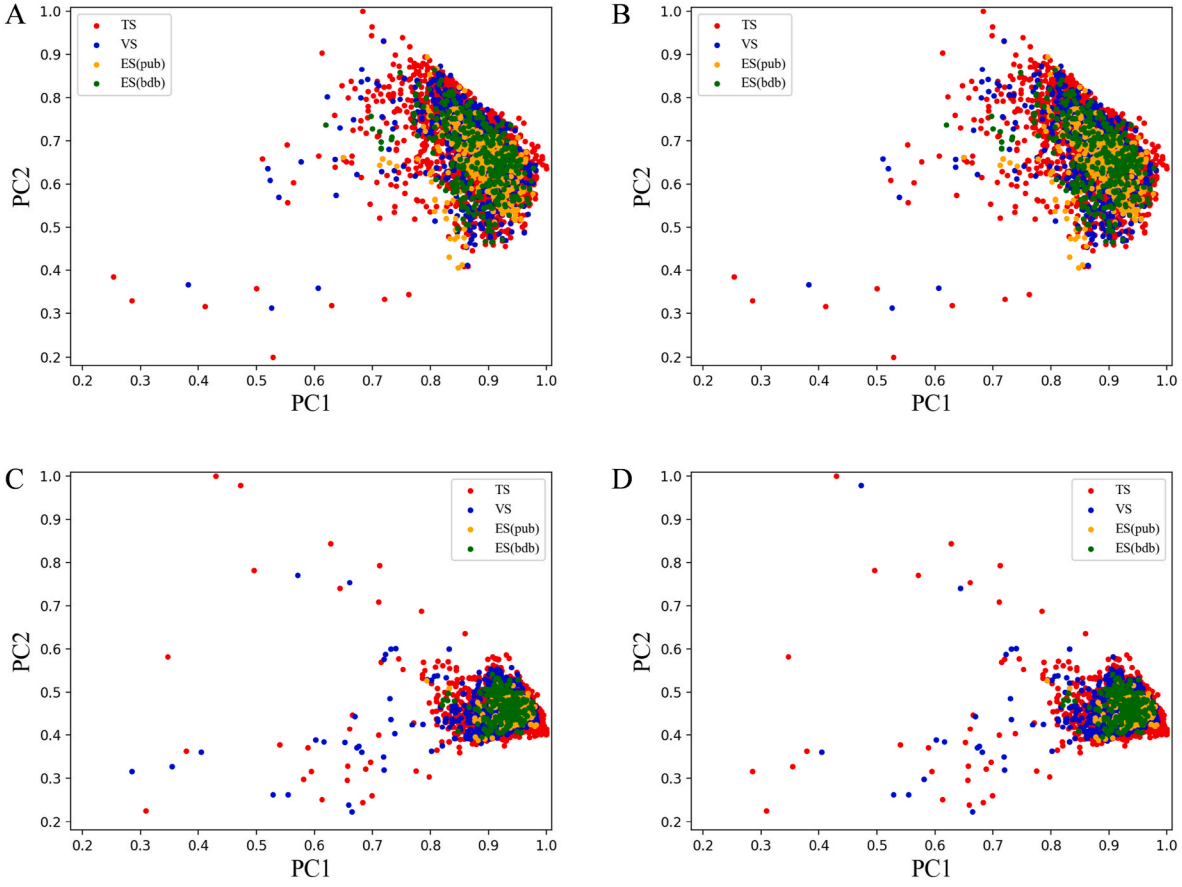

Fig. 2. PCA based on the physicochemical properties returned by the compounds belonging to TS, VS, and ES: A) CB2R affinity classifiers ($pK_i = 6.5$ as affinity threshold); B) CB2R affinity classifiers ($pK_i = 7$); C) CB1R affinity classifiers ($pK_i = 5.5$) and D) CB1R affinity classifiers ($pK_i = 6$).

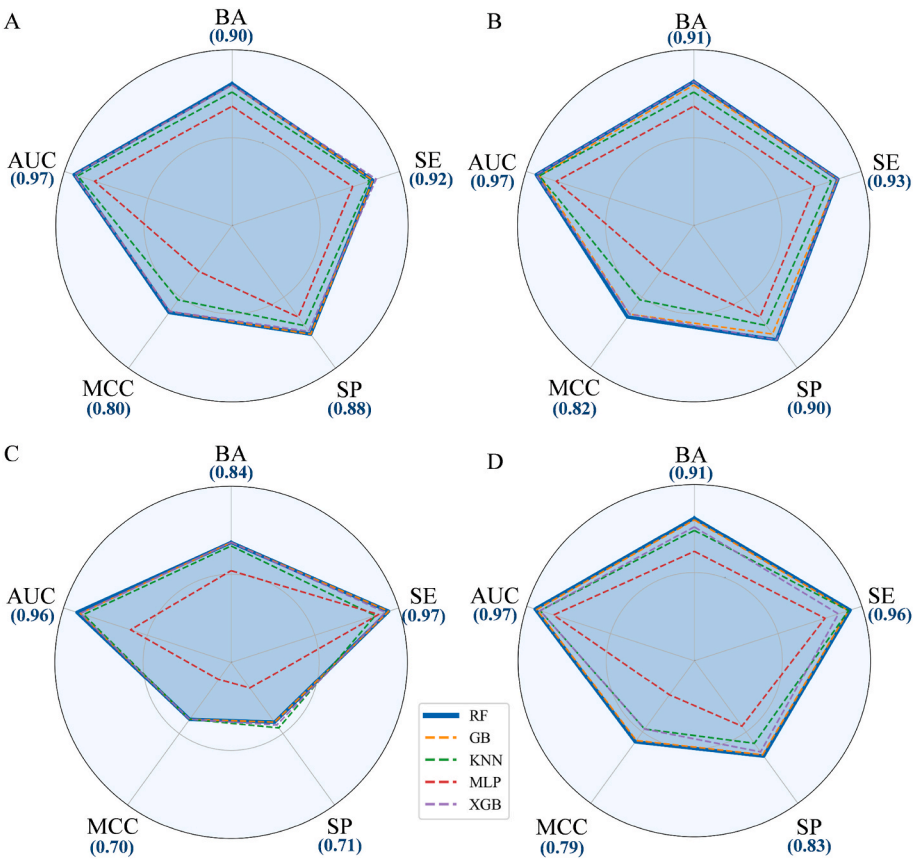

Fig. 3. Radar plots comparing the performance of the models on the VS using pK_i thresholds of A) 6.5 for CB2R; B) 7.0 for CB2R; C) 5.5 for CB1R and D) 6.0 for CB1R. For each model, the following statistics are reported: Balanced Accuracy (BA), Sensitivity (SE), Specificity (SP), Matthews Correlation Coefficient (MCC) and Area Under the Curve (AUC). The top-performing model (RF) is represented by a solid line, and its statistics are reported within parentheses. All the others are depicted by dotted lines.

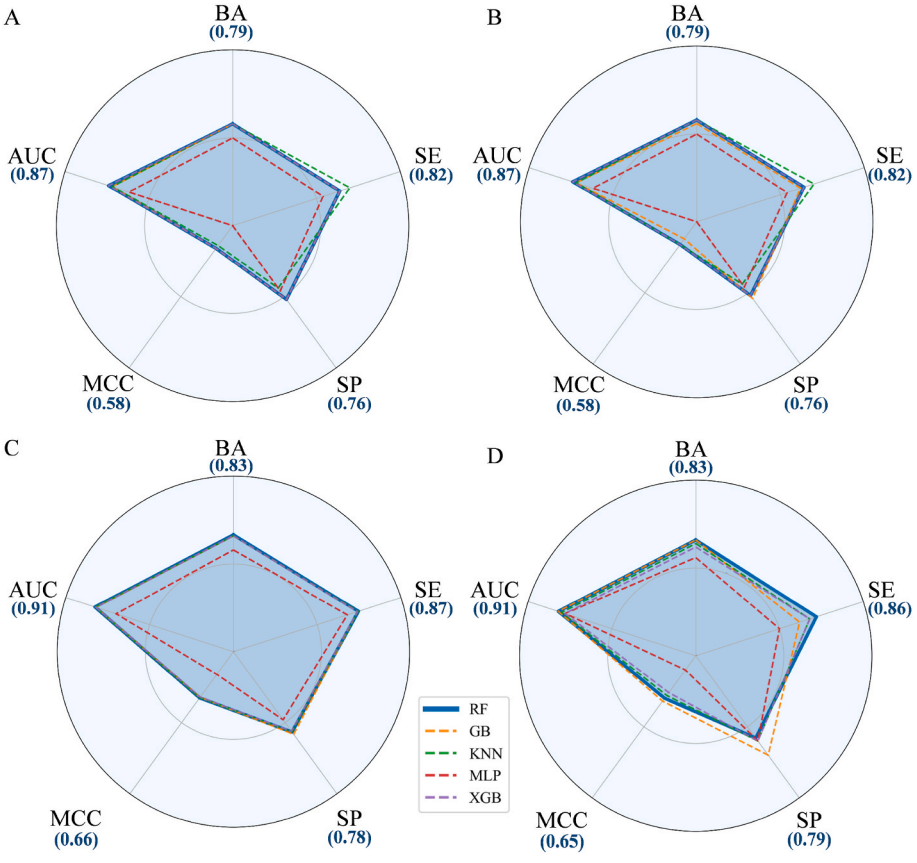

Fig. 4. Radar plots comparing the performance of the models in 5-fold cross-validation (5-CV) using pK_i thresholds of A) 6.5 for CB2R; B) 7.0 for CB2R; C) 5.5 for CB1R and D) 6.0 for CB1R. For each model, the following statistics are reported: Balanced Accuracy (BA), Sensitivity (SE), Specificity (SP), Matthews Correlation Coefficient (MCC) and Area Under the Curve (AUC). The top-performing model (RF) is represented by a solid line, and its statistics are reported within parentheses. All the others are depicted by dotted lines.

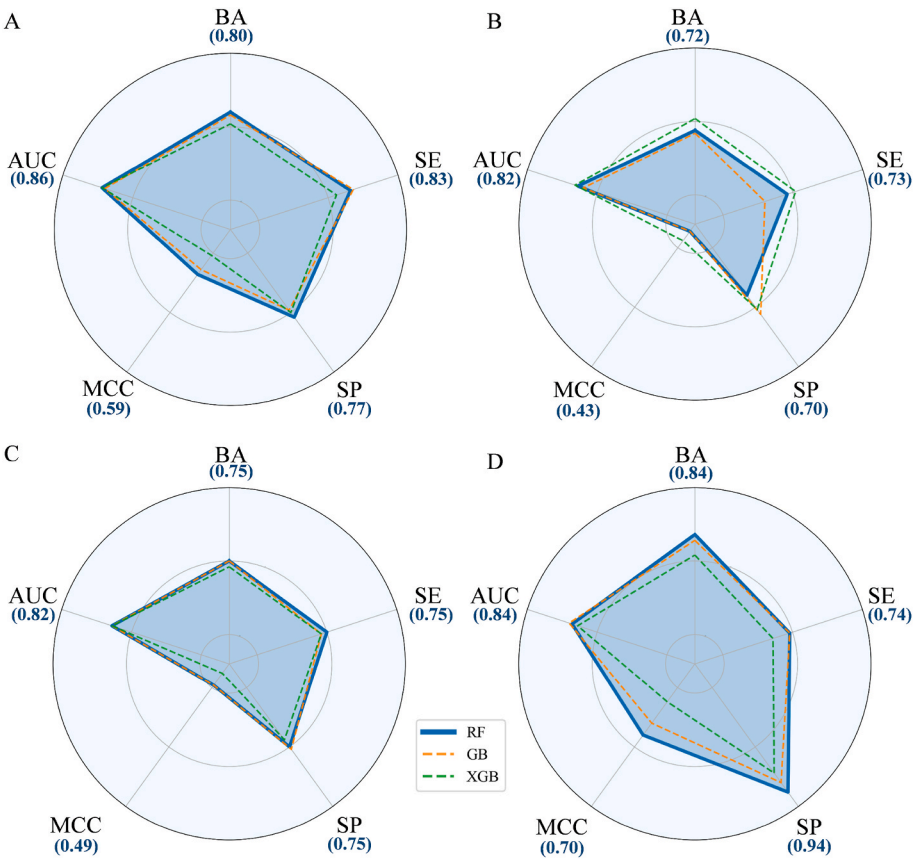

Fig. 5. Radar plots comparing the performance of the models on the ES extracted by Pubchem (ES(pub)) using pK_i thresholds of A) 6.5 for CB2R; B) 7.0 for CB2R; C) 5.5 for CB1R and D) 6.0 for CB1R. For each model, the following statistics are reported: Balanced Accuracy (BA), Sensitivity (SE), Specificity (SP), Matthews Correlation Coefficient (MCC) and Area Under the Curve (AUC). The top-performing model (RF) is represented by a solid line, and its statistics are reported within parentheses. All the others are depicted by dotted lines.

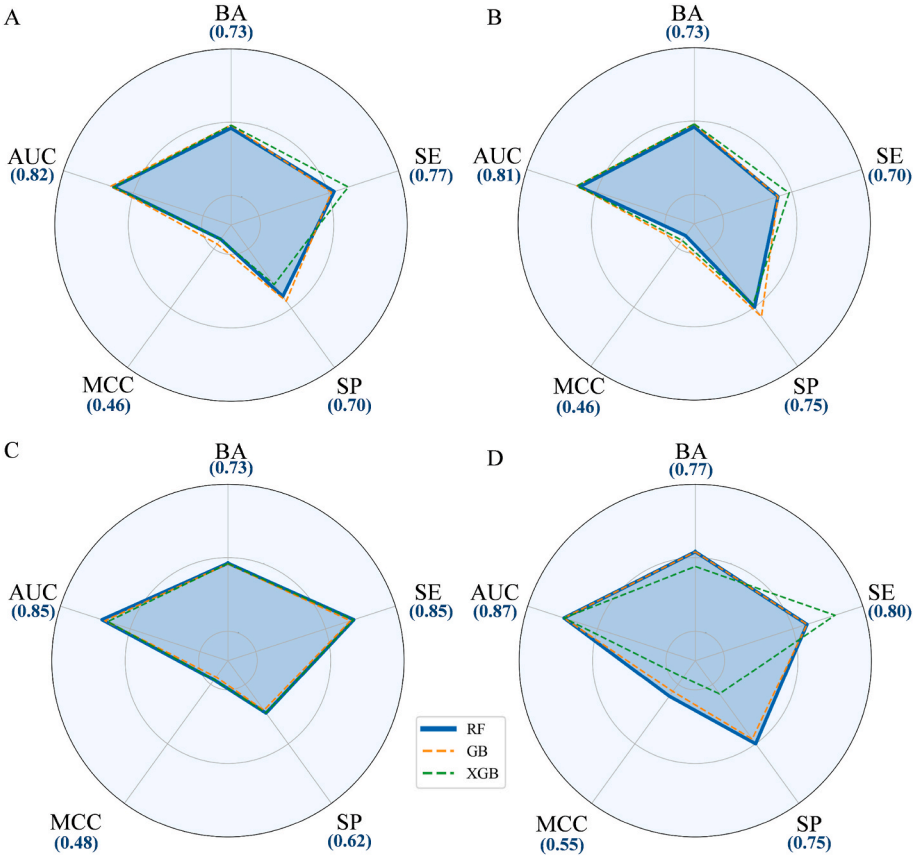

Fig. 6. Radar plots comparing the performance of the models on the ES extracted by BindingDB (ES(bdb) using pK_i thresholds of A) 6.5 for CB2R; B) 7.0 for CB2R; C) 5.5 for CB1R and D) 6.0 for CB1R. For each model, the following statistics are reported: Balanced Accuracy (BA), Sensitivity (SE), Specificity (SP), Matthews Correlation Coefficient (MCC) and Area Under the Curve (AUC). The top-performing model (RF) is represented by a solid line, and its statistics are reported within parentheses. All the others are depicted by dotted lines.

using two independent ES (Fig. 5B and 6B and Table S4) support the robustness and generalizability of these models. Indeed, BA values reach a maximum of 0.76 (XGB) and 0.74 (XGB) when *ES(pub)* and *ES(bdb)* are considered, respectively. Impressive AUC values were also observed, being all above 0.80 (Table S4).

3.2. Classification models of CB1R affinity

3.2.1. $pK_i = 5.5$ as affinity threshold

Fig. 3C reports the performances in validation obtained from the developed CB1R affinity classifiers using as affinity threshold $pK_i = 5.5$ (see Table S2 for reading the values of the quality metrics returned by all the models**)**. It is important to note that all the models were trained using a dataset where the number of LAF examples slightly exceeded that of HAF examples, as indicated in Table 1 (IR = 1.06 for both TS and VS). Once again, RF, GB, and XGB emerged as the top-performing algorithms, displaying high BA (0.84), MCC (0.70) and AUC ($\geq$0.95) values during the validation process. Notably, these classifiers maintain their high predictive power during the performed 5-fold cross-validation (BA = 0.83 and AUC = 0.91), as reported in Fig. 4C and Table S3. Satisfactory performances were also achieved in external validation, with AUC values ranging from 0.82 to 0.85 and BA values ranging from 0.73 to 0.75 ((Fig. 5C and 6C and Table S4).

3.2.2. $pK_i = 6$ as affinity threshold

For the CB1R classifiers developed using a pK_i of 6 as affinity threshold, the findings discussed earlier were further confirmed. Firstly, RF, GB, and XGB are the top-performing algorithms, as responsible for high values of BA ranging from 0.88 to 0.90, MCC ranging from 0.74 to 0.78, and an AUC of 0.97 during the validation process (Fig. 3D and Table S2). Secondly, these models maintained their predictive power after undergoing 5-CV, as indicated by the BA value of 0.83 and AUC of 0.91 (Fig. 4D and Table S3). Thirdly, the models demonstrated the ability to accurately predict compounds from the two considered ES, as reflected in the BA values ranging from 0.77 to 0.82 and AUC values from 0.84 to 0.87 (Fig. 5D and 6D and Table S4).

3.3. Algorithm selection

Taken as a whole, all the discussed data put forward the RF, GB and XGB algorithms as those of choice for obtaining reliable CB1R and CB2R affinity predictions. Remarkably, this finding is confirmed irrespective of the considered affinity threshold and receptor. Additionally, we developed a consensus model that utilizes the majority voting rule to assign a sample to the most frequent class. However, it is noteworthy that the results obtained from this model do not exhibit significant improvements compared to those obtained from a single model, as depicted in Table S5. As a result, we chose to focus on selecting the best model rather than combining them through consensus. To make the final selection, we prioritized the area under the curve (AUC), which indicates the ability of the models to differentiate between HAF and LAF. Based on obtained results, the RF algorithm appears the most promising for predicting CB1R and CB2R affinities. It displayed the highest AUC values for both receptors across all affinity thresholds and validation procedures (AUC ranging from 0.96 to 0.97 in validation and from 0.87 to 0.91 in 5-CV for CB2R and CB1R respectively). Encouraged by these results, we selected RF-based classifiers for a successive evaluation of performance stability.

3.4. Stability evaluation

In order to assess the stability of the selected RF-based models, an additional analysis was conducted. Specifically, for each model, 100 additional classifiers were constructed using different randomly selected TS and VS. It is worth noting that the data splitting was performed according to the methodology outlined in the Materials and Methods section. For each additional model, BA, SE, SP, AUC and MCC were computed, and the relative standard deviation evaluated to gauge the stability of these classifiers. Fig. 7A illustrates the average BA, AUC, SE, SP, and MCC values, along with their corresponding standard deviations.

It is evident that all the classifiers exhibit robustness and independence from the composition of the TS and VS. Notably, the computed standard deviations were consistently low ($\leq$0.02). Additionally, the box plot in Fig. 7B–D demonstrates a symmetrical distribution of performance metrics across the 100 models, with no significant outliers. The interquartile ranges (IQR) ranged from 0.01 to 0.03, indicating a narrow distribution of the obtained performance metrics around the mean. Overall, the stability assessment provides compelling evidence that the proposed models are robust and reliable, as they do not suffer from overfitting to the TS and be well-generalized to unseen data.

3.5. Classification model of CB2R/CB1R selectivity

Based on the developed models, a compound can be predicted to have high affinity towards CB2R ($pK_i > 7$ or 6.5) and significant affinity towards CB1R ($pK_i > 6$ or 5.5). In such cases, additional information on CB2R selectivity becomes valuable. To address this, we developed an RF-based model capable of directly predicting CB2R/CB1R selectivity rather than affinity towards a specific receptor only. The model was built using the *CBRs-qsarSEL* dataset, which includes **2183** compounds along with their relative $pK_{i(CB2R-CB1R)}$ values. Table 2 presents the quality metrics obtained from the model through 5-fold cross-validation, validation and external validation using an evaluation set consisting of 129 compounds (*ES_CBRs*), both selective and non-selective (detailed in the Materials and Methods section).

Notably, the classifier handles three classes of compounds: non-selective (*no-sel*), CB2R selective (*CB2R-sel*), and CB1R selective (*CB1R-sel)*. Remarkably, the model demonstrated satisfactory and balanced performance in predicting all the three classes during validation. It achieved BA and MCC values above 0.80 and 0.65, respectively, with consistently high AUC values exceeding 0.90. The model also exhibited stability in 5-CV, with an average BA and MCC of 0.80 and 0.60, respectively. Importantly, the reliability and generalizability of the model were confirmed during external validation, where BA ranged from 0.70 for the prediction of the *no-sel* class (non-selective compounds) to 0.87 for the *CB1R-sel* class (CB1R selective compounds). Importantly, AUC values ranged from 0.74 (*no-sel* class) to 0.87 (*CB1R-sel* class), further supporting the model ability to distinguish between the three classes.

3.6. ALPACA: a freely accessible web platform

The high-performing classifiers, developed using the Random Forest (RF) algorithm, are available for use on a freely accessible web-platform called ALPACA (A machine Learning Platform for Affinity and selectivity profiling of CAnnabinoids receptors modulators), accessible at https://www.ba.ic.cnr.it/softwareic/alpaca/. To use the platform, users have two options. They can either draw a 2D structure of their query molecule using the JSME canvas applet [59] or input a SMILES string directly into the provided text field. In addition, users can import a.txt file that contains a list of SMILES strings for virtual screening purposes by clicking on the "MASSIVE" button. Once the process is complete, the tool generates predictions concerning the affinity of the query towards CB2R and CB1R. For CB2R, the predictions are based on a threshold of $pK_i = 6.5$ and $pK_i = 7$, while for CB1R, the threshold is set at $pK_i = 5.5$ and $pK_i = 6$. The results are displayed as "YES" if the affinity exceeds the specified threshold and "NO" if it falls below. Furthermore, the tool provides information on the reliability of the predictions, taking into account the considered AD. If the query is predicted to be affine towards both CB2R and CB1R (above the respective thresholds), the tertiary classifier of CB2R/CB1R selectivity also generates a prediction. Users

A

	pK_i threshold	BA	SE	SP	MCC	AUC
CB2R	**6.5**	0.89±0.01	0.93±0.01	0.86±0.01	0.79±0.02	0.96±0.01
	7	0.89±0.01	0.91±0.01	0.89±0.01	0.79±0.02	0.96±0.01
CB1R	**5.5**	0.83±0.01	0.94±0.02	0.71±0.01	0.68±0.02	0.95±0.01
	6	0.90±0.01	0.97±0.01	0.83±0.01	0.78±0.01	0.96±0.01

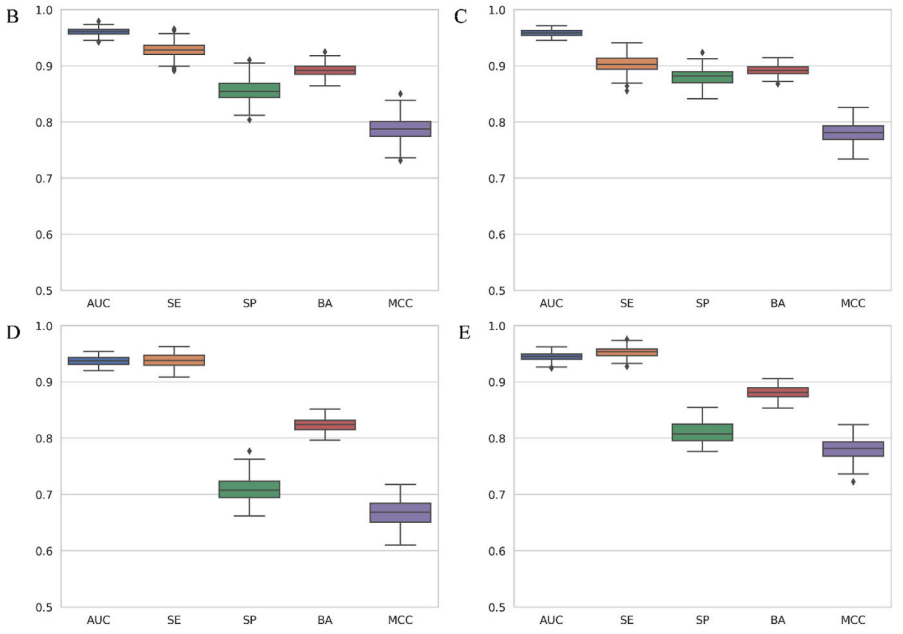

Fig. 7. A) Table showing the mean of each evaluated metric (BA, SP, SE, AUC, MCC) and the relative standard deviation returned by the performed stability evaluation applied to the classifiers developed using RF as algorithm; **B-C-D-E)** Box plot reporting the stability evaluation results. Each box represents the distribution of values for a given metric at the different pK_i threshold: **B)** 6.5 for CB2R; **C)** 7.0 for CB2R; **D)** 5.5 for CB1R and **E)** 6.0 for CB1R. The horizontal black line within each box represents the median value. The top and bottom of the coloured box represent the first and third quartiles, respectively. The black horizontal lines (whiskers), upper the box limits, represent observations within 1.5 times the Q1 and Q3. The point above the whiskers illustrates the outlier (observations above 1.5 times the Q1 and Q3).

Table 2
Quality metrics returned by the CB1R-CB2R_sel model on validation, 5-CV and external validation.

RF	Selectivity	BA	SE	SP	MCC	AUC	TP	FP	TN	FN
Validation	CB1R-sel	0.88	0.78	0.97	0.76	0.95	46	11	367	13
	CB2R-sel	0.84	0.88	0.81	0.69	0.90	183	43	185	26
	no-sel	0.83	0.75	0.90	0.66	0.90	127	27	241	42
5-CV	CB1R-sel	0.84	0.71	0.96	0.71	0.94	171	49	1465	69
	CB2R-sel	0.81	0.83	0.80	0.63	0.90	694	188	731	141
	no-sel	0.81	0.75	0.87	0.62	0.90	507	145	930	172
External Validation	CB1R-sel	0.87	0.76	0.98	0.79	0.87	12	2	106	5
	CB2R-sel	0.74	0.74	0.73	0.47	0.77	43	19	52	15
	no-sel	0.70	0.62	0.77	0.39	0.74	31	18	61	19

have the option to download the output as a.csv file. It is important to note that links to download the obtained data are sent to the user's registered email address. Additionally, the "History" page keeps a record of all user executions, including input SMILES and corresponding output. Fig. 8 displays the output page generated by the tool when using a CB2R selective ligand, published in April 2023 [60], which is not present in ChEMBLv30.

4. Conclusions

Although highly desirable for treating cancer and neurodegeneration, designing selective CB2R ligands is extremely challenging due to the very high similarity with the CB1R binding site, whose modulation is responsible for severe psychotropic side-effects. To address this issue and provide a convenient tool for medicinal chemists working in this field, we developed a series of classifiers to predict *i)* CB2R affinity, *ii)* CB1R affinity, and *iii)* CB2R/CB1R selectivity. We applied the best data-curation practices on bioactivity data reported in ChEMBLv30 on human CB1R and human CB2R and employed five ML algorithms (RF, KNN, GB, XGB and MLP) to develop, based on AP fingerprints, binary models differing for the considered affinity threshold. The analysis of the obtained performances in both internal and external validation put forward the RF, GB and XGB algorithms as those of choice for obtaining reliable CB1R and CB2R affinity predictions (BA and MCC in validation ranging from 0.84 to 0.91 and from 0.70 to 0.82, respectively). Notably, this finding held true regardless of the affinity threshold or receptor under consideration. Among those, RF was selected as the top-performing algorithm being responsible for the highest AUC values in validation ($\geq$0.96) and for this reason employed to build a tertiary classifier of CB2R/CB1R selectivity displaying satisfactory performances (BA $\geq$ 0.83 and AUC $\geq$0.90). Altogether, these encouraging results prompted us to make available the top-performing classifiers in a user-friendly web platform developed by our group and called ALPACA (https://www.ba.ic.cnr.it/softwareic/alpaca/). ALPACA, which can be used in combination with structure-based strategies, offers the capability to conduct virtual screening campaigns by inputting a compound library, enabling the identification of highly probable CB2R selective molecules with a high level of confidence. It is worth noting that our approach differs from those employed in similar papers [26,28] in two key aspects. Firstly, we developed classifiers using specific thresholds instead of regression models. Secondly, we have made the developed predictive tools available on a freely accessible

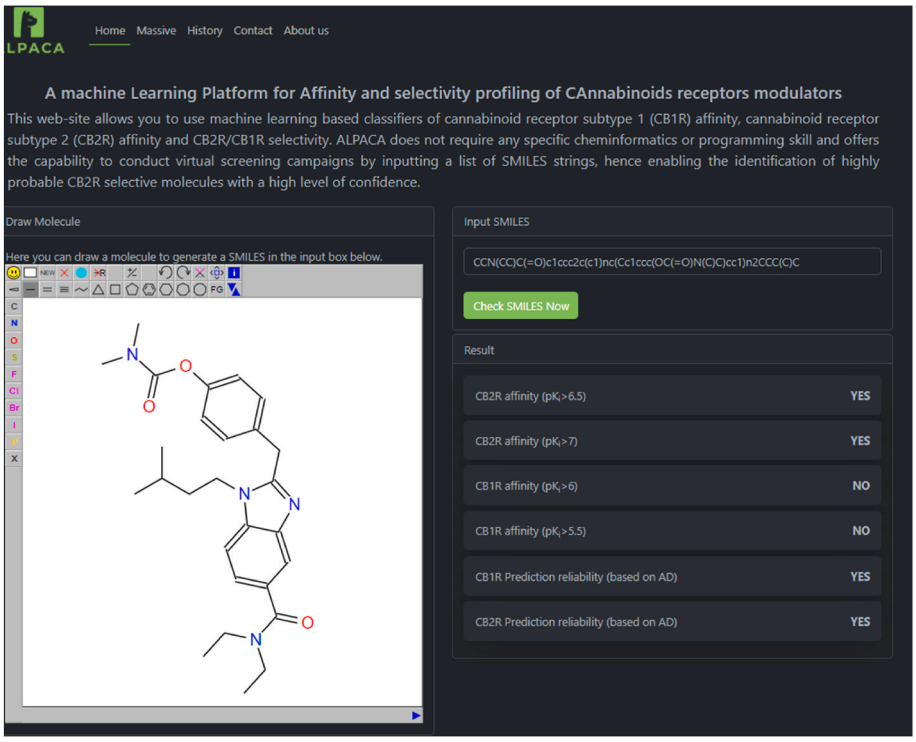

Fig. 8. Output page returned by the ALPACA web platform after using as query the SMILES string of a recently published [60] CB2R selective ligand.

web-platform, eliminating the need for specialized cheminformatics or programming skills. Based on the discussed data, we firmly believe that ALPACA can serve as a valuable tool for medicinal chemists interested in designing selective CB2R modulators for the treatment of cancer and neurodegenerative conditions.

Data and software availability

The following files are made available in the supporting information.

- *CB2R-qsarDB.xlsx* Excel file containing the **3514** SMILES strings of the chemicals belonging to the CB2R-qsarDB dataset and the corresponding experimental values (pK_i mean).*
- *CB1R-qsarDB.xlsx* Excel file containing the **3846** SMILES strings of the chemicals belonging to the CB1R-qsarDB dataset and the corresponding experimental values (pK_i mean).*
- *CBRs-qsarSEL.xslx* Excel file containing the **2183** SMILES strings of the chemicals belonging to the CBRs-qsarSEL dataset and the corresponding experimental values for CB2R and CB1R (pK_i mean).
- *ES(pub)_CB2R.xslx* Excel file containing the **373** SMILES strings of the chemicals belonging to the ES(pub)_CB2R dataset and the corresponding experimental values (pK_i mean).*
- *ES(bdb)_CB2R.xslx* Excel file containing the **355** SMILES strings of the chemicals belonging to the ES(bdb)_CB2R dataset and the corresponding experimental values (pK_i mean).*
- *ES(pub)_CB1R.xslx* Excel file containing the **238** SMILES strings of the chemicals belonging to the ES(pub)_CB1R dataset and the corresponding experimental values (pK_i mean).*
- *ES(bdb)_CB1R.xslx* Excel file containing the **268** SMILES strings of the chemicals belonging to the ES(bdb)_CB1R and the corresponding experimental values (pK_i mean).*
- *ES_CBRs.xslx* Excel file containing the **129** SMILES strings of the chemicals belonging to the ES_CBRs dataset and the corresponding experimental values for CB2R and CB1R (pK_i mean).

Please notice that within these files * **NA** in the field of pK_i indicates the absence of any affinity.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Dr. Ivan Mercurio for providing graphical support in creating the ALPACA logo and Dr. Candida Giovannelli for technical support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2023.107314.

References

[1] M. Contino, P.J. McCormick, Editorial: the canonical and non-canonical endocannabinoid system as a target in cancer and acute and chronic pain, Front. Pharmacol. (2020) 11. https://www.frontiersin.org/articles/10.3389/fphar.2020.00312. (Accessed 5 June 2023).
[2] V. Di Marzo, F. Piscitelli, The endocannabinoid system and its modulation by phytocannabinoids, Neurotherapeutics 12 (2015) 692–698, https://doi.org/10.1007/s13311-015-0374-6.
[3] F. Spinelli, E. Capparelli, C. Abate, N.A. Colabufo, M. Contino, Perspectives of cannabinoid type 2 receptor (CB2R) ligands in neurodegenerative disorders: structure-affinity relationship (SAfiR) and structure-activity relationship (SAR) studies, J. Med. Chem. 60 (2017) 9913–9931, https://doi.org/10.1021/acs.jmedchem.7b00155.
[4] L. Cristino, T. Bisogno, V. Di Marzo, Cannabinoids and the expanded endocannabinoid system in neurological disorders, Nat. Rev. Neurol. 16 (2020) 9–29, https://doi.org/10.1038/s41582-019-0284-z.
[5] N. Mangal, S. Erridge, N. Habib, A. Sadanandam, V. Reebye, M.H. Sodergren, Cannabinoids in the landscape of cancer, J. Cancer Res. Clin. Oncol. 147 (2021) 2507–2534, https://doi.org/10.1007/s00432-021-03710-7.

[6] C. Turcotte, M.-R. Blanchet, M. Laviolette, N. Flamand, The CB2 receptor and its role as a regulator of inflammation, Cell. Mol. Life Sci. 73 (2016) 4449–4470, https://doi.org/10.1007/s00018-016-2300-4.

[7] J.A. Komorowska-Müller, A.-C. Schmöle, CB2 receptor in microglia: the guardian of self-control, Int. J. Mol. Sci. 22 (2021) 19, https://doi.org/10.3390/ijms22010019.

[8] M. Tanaka, S. Sackett, Y. Zhang, Endocannabinoid modulation of microglial phenotypes in neuropathology, Front. Neurol. 11 (2020). https://www.frontiersin.org/articles/10.3389/fneur.2020.00087. (Accessed 5 June 2023).

[9] M.F. Nagoor Meeran, C. Sharma, S.N. Goyal, S. Kumar, S. Ojha, CB2 receptor-selective agonists as candidates for targeting infection, inflammation, and immunity in SARS-CoV-2 infections, Drug Dev. Res. 82 (2021) 7–11, https://doi.org/10.1002/ddr.21752.

[10] M. Rastegar, S. Samadizadeh, M. Yasaghi, A. Moradi, A. Tabarraei, V. Salimi, A. Tahamtan, Functional variation (Q63R) in the cannabinoid CB2 receptor may affect the severity of COVID-19: a human study and molecular docking, Arch. Virol. 166 (2021) 3117–3126, https://doi.org/10.1007/s00705-021-05223-7.

[11] A.C. Howlett, M.E. Abood, CB1 & CB2 receptor pharmacology, Adv. Pharmacol. 80 (2017) 169–206, https://doi.org/10.1016/bs.apha.2017.03.007.

[12] X. Li, T. Hua, K. Vemuri, J.-H. Ho, Y. Wu, L. Wu, P. Popov, O. Benchama, N. Zvonok, K. Locke, L. Qu, G.W. Han, M.R. Iyer, R. Cinar, N.J. Coffey, J. Wang, M. Wu, V. Katritch, S. Zhao, G. Kunos, L.M. Bohn, A. Makriyannis, R.C. Stevens, Z.-J. Liu, Crystal structure of the human cannabinoid receptor CB2, Cell 176 (2019) 459–467.e13, https://doi.org/10.1016/j.cell.2018.12.011.

[13] T. Hua, X. Li, L. Wu, C. Iliopoulos-Tsoutsouvas, Y. Wang, M. Wu, L. Shen, C. A. Brust, S.P. Nikas, F. Song, X. Song, S. Yuan, Q. Sun, Y. Wu, S. Jiang, T.W. Grim, O. Benchama, E.L. Stahl, N. Zvonok, S. Zhao, L.M. Bohn, A. Makriyannis, Z.-J. Liu, Activation and signaling mechanism revealed by cannabinoid receptor-gi complex structures, Cell 180 (2020) 655–665.e18, https://doi.org/10.1016/j.cell.2020.01.008.

[14] G. Graziano, P. Delre, F. Carofiglio, J. Brea, A. Ligresti, M. Kostrzewa, C. Riganti, C. Gioè-Gallo, M. Majellaro, O. Nicolotti, N.A. Colabufo, C. Abate, M.I. Loza, E. Sotelo, G.F. Mangiatordi, M. Contino, A. Stefanachi, F. Leonetti, N-adamantyl-anthranil amide derivatives: new selective ligands for the cannabinoid receptor subtype 2 (CB2R), Eur. J. Med. Chem. 248 (2023), 115109, https://doi.org/10.1016/j.ejmech.2023.115109.

[15] F. Intranuovo, L. Brunetti, P. DelRe, G.F. Mangiatordi, A. Stefanachi, A. Laghezza, M. Niso, F. Leonetti, F. Loiodice, A. Ligresti, M. Kostrzewa, J. Brea, M.I. Loza, E. Sotelo, M. Saviano, N.A. Colabufo, C. Riganti, C. Abate, M. Contino, Development of N-(1-Adamantyl)benzamides as novel anti-inflammatory multitarget agents acting as dual modulators of the cannabinoid CB2 receptor and fatty acid amide hydrolase, J. Med. Chem. 66 (2023) 235–250, https://doi.org/10.1021/acs.jmedchem.2c01084.

[16] G.F. Mangiatordi, F. Intranuovo, P. Delre, F.S. Abatematteo, C. Abate, M. Niso, T. M. Creanza, N. Ancona, A. Stefanachi, M. Contino, Cannabinoid receptor subtype 2 (CB2R) in a multitarget approach: perspective of an innovative strategy in cancer and neurodegeneration, J. Med. Chem. 63 (2020) 14448–14469, https://doi.org/10.1021/acs.jmedchem.0c01357.

[17] J.-F. Yang, A.H. Williams, N.R. Penthala, P.L. Prather, P.A. Crooks, C.-G. Zhan, Binding modes and selectivity of cannabinoid 1 (CB1) and cannabinoid 2 (CB2) receptor ligands, ACS Chem. Neurosci. 11 (2020) 3455–3463, https://doi.org/10.1021/acschemneuro.0c00551.

[18] G. Tu, T. Fu, F. Yang, J. Yang, Z. Zhang, X. Yao, W. Xue, F. Zhu, Understanding the polypharmacological profiles of triple reuptake inhibitors by molecular simulation, ACS Chem. Neurosci. 12 (2021) 2013–2026, https://doi.org/10.1021/acschemneuro.1c00127.

[19] G. Zheng, F. Yang, T. Fu, G. Tu, Y. Chen, X. Yao, W. Xue, F. Zhu, Computational characterization of the selective inhibition of human norepinephrine and serotonin transporters by an escitalopram scaffold, Phys. Chem. Chem. Phys. 20 (2018) 29513–29527, https://doi.org/10.1039/c8cp06232c.

[20] W. Xue, F. Yang, P. Wang, G. Zheng, Y. Chen, X. Yao, F. Zhu, What contributes to serotonin-norepinephrine reuptake inhibitors' dual-targeting mechanism? The key role of transmembrane domain 6 in human serotonin and norepinephrine transporters revealed by molecular dynamics simulation, ACS Chem. Neurosci. 9 (2018) 1128–1140, https://doi.org/10.1021/acschemneuro.7b00490.

[21] W. Xue, P. Wang, G. Tu, F. Yang, G. Zheng, X. Li, X. Li, Y. Chen, X. Yao, F. Zhu, Computational identification of the binding mechanism of a triple reuptake inhibitor amitifadine for the treatment of major depressive disorder, Phys. Chem. Chem. Phys. 20 (2018) 6606–6616, https://doi.org/10.1039/c7cp07869b.

[22] T. Fu, G. Zheng, G. Tu, F. Yang, Y. Chen, X. Yao, X. Li, W. Xue, F. Zhu, Exploring the binding mechanism of metabotropic glutamate receptor 5 negative allosteric modulators in clinical trials by molecular dynamics simulations, ACS Chem. Neurosci. 9 (2018) 1492–1502, https://doi.org/10.1021/acschemneuro.8b00059.

[23] A. Gaulton, L.J. Bellis, A.P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, J.P. Overington, ChEMBL: a large-scale bioactivity database for drug discovery, Nucleic Acids Res. 40 (2012) D1100–D1107, https://doi.org/10.1093/nar/gkr777.

[24] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B.A. Shoemaker, J. Wang, B. Yu, J. Zhang, S.H. Bryant, PubChem substance and compound databases, Nucleic Acids Res. 44 (2016) D1202–D1213, https://doi.org/10.1093/nar/gkv951.

[25] Y. Bian, Y. Jing, L. Wang, S. Ma, J.J. Jun, X.-Q. Xie, Prediction of orthosteric and allosteric regulations on cannabinoid receptors using supervised machine learning classifiers, Mol. Pharm. 16 (2019) 2605–2615, https://doi.org/10.1021/acs.molpharmaceut.9b00182.

[26] H. Zhou, M. Shan, L.-P. Qin, G. Cheng, Reliable prediction of cannabinoid receptor 2 ligand by machine learning based on combined fingerprints, Comput. Biol. Med. 152 (2023), 106379, https://doi.org/10.1016/j.compbiomed.2022.106379.

[27] D. Ruano-Ordás, L. Burggraaff, R. Liu, C. van der Horst, L.H. Heitman, M.T. M. Emmerich, J.R. Mendez, I. Yevseyeva, G.J.P. van Westen, A multiple classifier system identifies novel cannabinoid CB2 receptor ligands, J. Cheminf. 11 (2019) 66, https://doi.org/10.1186/s13321-019-0389-9.

[28] M. Mizera, D. Latek, J. Cielecka-Piontek, Virtual screening of C. Sativa constituents for the identification of selective ligands for cannabinoid receptor 2, Int. J. Mol. Sci. 21 (2020) 5308, https://doi.org/10.3390/ijms21155308.

[29] P. Delre, G.J. Lavado, G. Lamanna, M. Saviano, A. Roncaglioni, E. Benfenati, G. F. Mangiatordi, D. Gadaleta, Ligand-based prediction of hERG-mediated cardiotoxicity based on the integration of different machine learning techniques, Front. Pharmacol. (2022) 13. https://www.frontiersin.org/articles/10.3389/fphar.2022.951083. (Accessed 5 June 2023).

[30] T.M. Creanza, P. Delre, N. Ancona, G. Lentini, M. Saviano, G.F. Mangiatordi, Structure-based prediction of hERG-related cardiotoxicity: a benchmark study, J. Chem. Inf. Model. 61 (2021) 4758–4770, https://doi.org/10.1021/acs.jcim.1c00744.

[31] T.M. Creanza, G. Lamanna, P. Delre, M. Contino, N. Corriero, M. Saviano, G. F. Mangiatordi, N. Ancona, DeLA-Drug: a deep learning algorithm for automated design of druglike analogues, J. Chem. Inf. Model. 62 (2022) 1411–1424, https://doi.org/10.1021/acs.jcim.2c00205.

[32] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, Open Babel: an open chemical toolbox, J. Cheminf. 3 (2011) 33, https://doi.org/10.1186/1758-2946-3-33.

[33] M.R. Berthold, N. Cebron, F. Dill, T.R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, KNIME: the konstanz information miner, in: C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker (Eds.), Data Analysis, Machine Learning and Applications, Springer, Berlin, Heidelberg, 2008, pp. 319–326, https://doi.org/10.1007/978-3-540-78246-9_38.

[34] R.E. Carhart, D.H. Smith, R. Venkataraghavan, Atom pairs as molecular features in structure-activity studies: definition and applications, J. Chem. Inf. Comput. Sci. 25 (1985) 64–73, https://doi.org/10.1021/ci00046a002.

[35] X. Chen, Y. Lin, M.K. Gilson, The binding database: overview and user's guide, Biopolymers 61 (2001) 127–141, https://doi.org/10.1002/1097-0282(2002)61:2<127::AID-BIP10076>3.0.CO;2-N.

[36] RDKit: open-source cheminformatics. https://www.rdkit.org (n.d.).

[37] D. Bajusz, A. Rácz, K. Héberger, Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? J. Cheminf. 7 (2015) 20, https://doi.org/10.1186/s13321-015-0069-3.

[38] D. Baptista, J. Correia, B. Pereira, M. Rocha, Evaluating molecular representations in machine learning models for drug response prediction and interpretability, Journal of Integrative Bioinformatics 19 (2022), https://doi.org/10.1515/jib-2022-0006.

[39] K. Roy, S. Kar, R.N. Das, Chapter 7 - validation of QSAR models, in: K. Roy, S. Kar, R.N. Das (Eds.), Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, Academic Press, Boston, 2015, pp. 231–289, https://doi.org/10.1016/B978-0-12-801505-6.00007-7.

[40] O. Nicolotti, D. Gadaleta, G. Mangiatordi, M. Catto, A. Carotti, Applicability domain for QSAR models, International Journal of Quantitative Structure-Property Relationships 1 (2016) 45–63, https://doi.org/10.4018/IJQSPR.2016010102.

[41] R. Zhu, Y. Guo, J.-H. Xue, Adjusting the imbalance ratio by the dimensionality of imbalanced data, Pattern Recogn. Lett. 133 (2020) 217–223, https://doi.org/10.1016/j.patrec.2020.03.004.

[42] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philos Trans A Math Phys Eng Sci 374 (2016), 20150202, https://doi.org/10.1098/rsta.2015.0202.

[43] E.L. Willighagen, J.W. Mayfield, J. Alvarsson, A. Berg, L. Carlsson, N. Jeliazkova, S. Kuhn, T. Pluskal, M. Rojas-Chertó, O. Spjuth, G. Torrance, C.T. Evelo, R. Guha, C. Steinbeck, The Chemistry Development Kit (CDK) v2.0: atom typing, depiction, molecular formulas, and substructure searching, J. Cheminf. 9 (2017) 33, https://doi.org/10.1186/s13321-017-0220-4.

[44] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[45] N.S. Altman, An introduction to kernel and nearest-neighbor nonparametric regression, Am. Statistician 46 (1992) 175–185, https://doi.org/10.1080/00031305.1992.10475879.

[46] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (2001) 1189–1232, https://doi.org/10.1214/aos/1013203451.

[47] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA, 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.

[48] M. Kubat, Neural networks: a comprehensive foundation by simon haykin, macmillan, 1994, ISBN 0-02-352781-7, Knowl. Eng. Rev. 13 (1999) 409–412, https://doi.org/10.1017/S0269888998214044.

[49] P. Refaeilzadeh, L. Tang, H. Liu, Cross-validation, in: L. LIU, M.T. ÖZSU (Eds.), Encyclopedia of Database Systems, Springer US, Boston, MA, 2009, pp. 532–538, https://doi.org/10.1007/978-0-387-39940-9_565.

[50] S.M. LaValle, M.S. Branicky, On the relationship between classical grid search and probabilistic roadmaps, in: J.-D. Boissonnat, J. Burdick, K. Goldberg, S. Hutchinson (Eds.), Algorithmic Foundations of Robotics V, Springer, Berlin, Heidelberg, 2004, pp. 59–75, https://doi.org/10.1007/978-3-540-45058-0_5.

[51] J. Snoek, H. Larochelle, R.P. Adams, Practical bayesian optimization of machine learning algorithms. http://arxiv.org/abs/1206.2944, 2012. (Accessed 5 June 2023).
[52] R. Trevethan, Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice, Front. Public Health (2017) 5. https://www.frontiersin.org/articles/10.3389/fpubh.2017.00307. (Accessed 5 June 2023).
[53] D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC Genom. 21 (2020) 6, https://doi.org/10.1186/s12864-019-6413-7.
[54] F. Melo, Area under the ROC curve, in: W. Dubitzky, O. Wolkenhauer, K.-H. Cho, H. Yokota (Eds.), Encyclopedia of Systems Biology, Springer, New York, NY, 2013, pp. 38–39, https://doi.org/10.1007/978-1-4419-9863-7_209.
[55] G. Melagraki, A. Afantitis, H. Sarimveis, O. Igglessi-Markopoulou, P. Koutentis, G. Kollias, In silico exploration for identifying structure-activity relationship of MEK inhibition and oral bioavailability for isothiazole derivatives, Chem. Biol. Drug Des. 76 (2010) 397–406, https://doi.org/10.1111/j.1747-0285.2010.01029.x.
[56] A. Afantitis, G. Melagraki, P.A. Koutentis, H. Sarimveis, G. Kollias, Ligand-based virtual screening procedure for the prediction and the identification of novel β-amyloid aggregation inhibitors using Kohonen maps and Counterpropagation Artificial Neural Networks, Eur. J. Med. Chem. 46 (2011) 497–508, https://doi.org/10.1016/j.ejmech.2010.11.029.
[57] S.Y. Ho, K. Phua, L. Wong, W.W. Bin Goh, Extensions of the external validation for checking learned model interpretability and generalizability, Patterns 1 (2020), 100129, https://doi.org/10.1016/j.patter.2020.100129.
[58] S. Fotouhi, S. Asadi, M.W. Kattan, A comprehensive data level analysis for cancer diagnosis on imbalanced data, J. Biomed. Inf. 90 (2019), 103089, https://doi.org/10.1016/j.jbi.2018.12.003.
[59] B. Bienfait, P. Ertl, JSME: a free molecule editor in JavaScript, J. Cheminf. 5 (2013) 24, https://doi.org/10.1186/1758-2946-5-24.
[60] P. Spatz, S.A.M. Steinmüller, A. Tutov, E. Poeta, A. Morilleau, A. Carles, M. H. Deventer, J. Hofmann, C.P. Stove, B. Monti, T. Maurice, M. Decker, Dual-acting small molecules: subtype-selective cannabinoid receptor 2 agonist/butyrylcholinesterase inhibitor hybrids show neuroprotection in an alzheimer's disease mouse model, J. Med. Chem. 66 (2023) 6414–6435, https://doi.org/10.1021/acs.jmedchem.3c00541.